

読み上げシステムおよび携帯端末装置

Reading System and Information Terminal

発明の属する技術分野

- 5 本発明は、テキストデータを音声変換する携帯端末装置および読み上げシステムに関する。

発明の背景

- 10 従来、電子メールやワープロの文章を読み上げる装置としては、記憶容量の豊富さや処理能力の高さ、ネットワーク機能の充実度等からパーソナルコンピュータにて実現していた。

- 15 しかしながらパーソナルコンピュータを歩きながら使用するには、大きさ、重量の問題から大変不便であるし、その操作も容易とは言い難い面がある。またテキストデータを音声変換するのみの機能であれば、コストパフォーマンスに欠ける等の問題がある。この点を解決するものとして、例えば特開平6-337774号公報には、情報処理装置への取り付け取り外しが簡単で、小型の情報処理装置（小型パーソナルコンピュータ等）にも内蔵でき、且つ小型軽量で持ち運びができると共に単体でも文章読み上げ機能を持つICカード形態の文章読み上げ装置が記載されている。このものは、予め文章読み上げ装置内のRAMにパーソナルコンピュータ等から外部インタフェースを介して文書データを転送しておき、
- 20 文章読み上げ装置が単体の時はRAMから文章データを読み出し、この読み出した文章データに対して言語処理を施して音声記号列を得、これを音声合成器に与えて音声データにし、この音声データをアナログ音声信号化してイヤホンジャックに出力する。この文章読み上げ装置はICカード化されているで、小型軽量且
- 25 つ着脱自在である。

また出力される音声も男性や女性といった一般的なものであり、必ずしもユーザが所望する声色での音声出力ではないので、ユーザが聴いていて楽しさを感じにくい面があった。

特開平7-140999号公報には、人間の発声に近い合成音声を生成することができる音声合成装置及び音声合成方法が開示されている。すなわち、辞書の中に読み仮名、アクセント型等の情報をととも、アクセント指令値及び又は音韻継続時間長情報を予め用意しておき、音韻の継続時間長を用いて音素片データのパラメータ列を生成するとともに、算出した音韻の継続時間長、韻律情報及びアクセント指令値に基づいてピッチパターンを生成し、このように生成されたパラメータ列及びピッチパターンにより音声波形を合成することにより、人間の発声に一段と近い合成音声を出力するものである。

また特開平11-143483号公報には、パソコン、ワープロ、ゲーム機、専用機、NC（ネットワークコンピュータ）、STB（セット・トップ・ボックス）などを利用する際の合成音声の発生に係わり、特にユーザが任意でかつ多様な合成音声を選ぶことが可能な手段を実現するシステムが開示されている。すなわち、人の音声を入力しその音声認識を行い、この認識した結果を解析し音韻系列作成のための音韻記号列情報をおよび韻律情報を抽出し、さらに特定の人の音声から作成した音声辞書（音声素片辞書）を準備しておき、前述の抽出した音韻記号列に基づいて音声素片を接続補間し音韻系列を作成する。

以上のように、音韻の継続時間長、韻律情報及びアクセント指令値に基づいて人間の発声に近い合成音声を出力するものも考えられているが、例えば文学作品を朗読させた場合、真に感動を与え、ユーザを楽しませるものとは限らない。

発明の概要

携帯性に優れ、歩行時においても無理無く持ち運べ、かつユーザが所望する声色で音声出力を行う端末装置と読上げシステムを提供する。その端末装置は音素をデータ化した音素データベースに基づき、音声合成目的データを音声合成処理する音声合成処理部を有する端末装置本体と、前記音声合成目的データと前記音素データベースとを保存し、前記端末装置本体と脱着可能な記憶装置とを備える。

図面の簡単な説明

図 1 は本発明の実施の形態 1 による携帯端末装置を有する通信システムを示す構成図である。

図 2 は実施の形態 1 による端末装置と記憶装置とを詳細に示すブロック図である。

5 図 3 は実施の形態 1 による端末装置の動作を示すフローチャートである。

図 4 は本発明の実施の形態 2 による携帯端末装置を有する通信システムを示す構成図である。

図 5 は実施の形態 1 による端末装置と記憶装置とを詳細に示すブロック図である。

10 図 6 は実施の形態 1 による端末装置の動作を示すフローチャートである。

図 7 は本発明の実施の形態 3 による携帯端末装置を有する通信システムを示す構成図である。

図 8 は実施の形態 3 による端末装置と記憶装置とを詳細に示すブロック図である。

15 図 9 は実施の形態 3 による端末装置の動作を示すフローチャートである。

図 10 は本発明の実施の形態 4 による携帯端末装置を有する通信システムを示す構成図である。

図 11 は実施の形態 4 による端末装置と記憶装置とを詳細に示すブロック図である。

20 図 12 は実施の形態 4 による端末装置の動作を示すフローチャートである。

図 13 は本発明の実施の形態 5 による携帯端末装置を有する通信システムを示す構成図である。

図 14 は実施の形態 5 による端末装置と記憶装置とを詳細に示すブロック図である。

25 図 15 は実施の形態 5 による端末装置の動作を示すフローチャートである。

図 16 は本発明の実施の形態 6 による携帯端末装置を有する通信システムを示す構成図である。

図 17 は実施の形態 6 による端末装置と記憶装置とを詳細に示すブロック図で

ある。

図 1 8 は実施の形態 6 による端末装置の動作を示すフローチャートである。

図 1 9 は本発明の実施の形態 7 による携帯端末装置を有する通信システムを示す構成図である。

- 5 図 2 0 は実施の形態 7 による端末装置と記憶装置とを詳細に示すブロック図である。

図 2 1 は実施の形態 7 による端末装置の動作を示すフローチャートである。

図 2 2 は本発明の実施の形態 8 による携帯端末装置を有する通信システムを示す構成図である。

- 10 図 2 3 は実施の形態 8 による端末装置と記憶装置とを詳細に示すブロック図である。

図 2 4 は実施の形態 8 による端末装置の動作を示すフローチャートである。

図 2 5 は本発明の実施の形態 9 による携帯端末装置を有する通信システムを示す構成図である。

- 15 図 2 6 は実施の形態 9 による端末装置と記憶装置とを詳細に示すブロック図である。

図 2 7 は本発明の実施の形態 1 0 による携帯端末装置を有する通信システムを示す構成図である。

- 20 図 2 8 は実施の形態 1 0 による端末装置と記憶装置とを詳細に示すブロック図である。

図 2 9 は本発明の実施の形態 1 1 による携帯端末装置を有する通信システムを示す構成図である。

図 3 0 は実施の形態 1 1 による端末装置と記憶装置とを詳細に示すブロック図である。

- 25 図 3 1 は実施の形態 1 1 による端末装置の動作を示すフローチャートである。

図 3 2 は本発明の実施の形態 1 2 による端末装置と記憶装置とを詳細に示すブロック図である。

図 3 3 は実施の形態 1 2 における携帯端末装置の翻訳処理部分のフローチャー

トである。

図 3 4 は本発明の実施の形態による携帯端末装置の詳細ブロック図である。

図 3 5 は本発明の実施の形態による携帯端末装置を有する通信システムを示す構成図及びサーバ装置内のブロック図である。

5 図 3 6 は実施の形態 1 4 におけるサーバー装置の動作フローチャートである。

図 3 7 は本発明の実施の形態による携帯端末装置を有する通信システムを示す構成図である。

図 3 8 は本発明の実施の形態による携帯端末装置と記憶装置とを詳細に示すブロック図である。

10 図 3 9 は実施の形態 1 5 における図 3 8 の携帯端末装置の動作フローチャートである。

図 4 0 は本発明の実施の形態による携帯端末装置のブロック図である。

図 4 1 は実施の形態 1 6 における図 4 0 の携帯端末装置の動作フローチャートである。

15 図 4 2 は実施の形態 1 7 における端末装置の動作フローチャートである。

図 4 3 は本発明の実施の形態による携帯端末装置を有する通信システムを示す構成図である。

図 4 4 は図 4 3 記載の携帯端末装置のブロック図である。

図 4 5 は図 4 4 記載の携帯端末装置の動作フローチャートである。

20 図 4 6 は本発明の実施の形態による携帯端末装置を有する通信システムを示す構成図である。

図 4 7 は本発明の実施の形態による携帯端末装置のブロック図である。

図 4 8 は本発明の実施の形態による携帯端末装置の動作フローチャートである。

図 4 9 は本発明の実施の形態による読上げシステムを示す構成図である。

25 図 5 0 は本発明の実施の形態による携帯端末装置とサーバー装置の詳細ブロック図である。

図 5 1 は本発明の実施の形態による読み上げシステムの動作フローチャートである。

図52は本発明の実施の形態による読上げシステムを示す構成図である。

図53は本発明の実施の形態による携帯端末装置とサーバー装置の詳細ブロック図である。

図54は本発明の実施の形態による読上げシステムの動作フローチャートである。

図55は本発明の実施の形態による読上げシステムを示す構成図である。

図56は本発明の実施の形態による携帯端末装置とサーバー装置の詳細ブロック図である。

図57は本発明の実施の形態による読上げシステムの動作フローチャートである。

図58は本発明の実施の形態による読上げシステムを示す構成図である。

図59は本発明の実施の形態による携帯端末装置とサーバー装置の詳細ブロック図である。

図60は本発明の実施の形態による読上げシステムを示す構成図である。

図61は本発明の実施の形態による携帯端末装置とサーバー装置の詳細ブロック図である。

図62は本発明の実施の形態による読上げシステムの動作フローチャートである。

図63は本発明の実施の形態による読上げシステムを示す構成図である。

図64は本発明の実施の形態による携帯端末装置とサーバー装置の詳細ブロック図である。

図65は本発明の実施の形態による読上げシステムの動作フローチャートである。

図66は本発明の実施の形態による読上げシステムを示す構成図である。

図67は本発明の実施の形態による携帯端末装置とサーバー装置の詳細ブロック図である。

図68は本発明の実施の形態による読上げシステムの動作フローチャートである。

図 6 9 は本発明の実施の形態による読み上げシステムの概略説明図である。

図 7 0 は本発明の実施の形態による携帯端末装置とサーバー装置の詳細ブロック図である。

図 7 1 は本発明の実施の形態による読み上げシステムの動作フローチャートである。

図 7 2 は本発明の実施の形態による読み上げシステムの概略説明図である。

図 7 3 は本発明の実施の形態による携帯端末装置とサーバー装置の詳細ブロック図である。

図 7 4 は本発明の実施の形態による読み上げシステムの動作フローチャート図である。

好ましい実施例の説明

(実施の形態 1)

図 1 は本発明の実施の形態 1 の携帯端末装置を有する通信システムを示す構成図である。実施の形態 1 および以下の各実施の形態において、携帯端末装置とは、端末装置と記憶装置とから構成される装置である。

図 1 において、端末装置 2 0 1 は表示部や操作部、ヘッドフォン・スピーカ等の音声出力部などを備える。メモリカード等の記憶装置 2 0 2 はテキストデータ等の音声合成目的データ 2 0 3 および音素データ 2 0 4 を記憶する。記憶装置 2 0 2 は端末装置 2 0 1 とは脱着可能である。インターネット上のサーバ装置 2 0 5 は音声合成目的データ 2 0 3 や音素データ 2 0 4、音声合成処理プログラム 2 0 6 を提供する。ここでは一つのサーバ装置しか記載していないが、複数のサーバ装置で音声合成目的データと音素データとを分けて提供する場合もある。ユーザは、キャラクタ音声の音素データと読み上げ対象の音声合成目的データとが格納された記憶装置 2 0 2 を端末装置 2 0 1 に挿入し操作することで、自分の所望のキャラクタ音声にて合成音声を聴くことができる。

図 2 は、図 1 の端末装置 2 0 1 と記憶装置 2 0 2 とを詳細に示すブロック図である。図 2 において、システム制御部 1 0 1 は装置 2 0 1 内の各処理部とデータ

のやり取りを行い、装置全体の制御を行う。音声合成処理部102は音声合成目的データの解析を行って各データに最適な音素データを抽出及び連結し、後述の音声出力処理部104に渡せるようにデータ変換する。記憶装置インタフェース(I/F)部103はシステム制御部101から指示を受け、記憶装置202に
5 対してデータの読み書きを行う。音声出力処理部104は音声合成処理部102からデータを受け、データのフォーマット変換を行い、スピーカまたはヘッドフォン108に出力する。記憶部105は装置201を制御するプログラムの保管やデータ処理の際の作業領域として用いられる。

ユーザは、操作部106を通じ、装置201に自分の指示を与える。表示部1
10 07は装置201の動作状態等をユーザに表示する。電源部109は装置201、202に電源を供給する。通信処理部110は公衆回線と接続して、インターネット上でのデータのやり取りを実現する。端末装置インタフェース(I/F)部120は記憶装置インタフェース部103と共に端末装置201とデータのやり取りを行う。音素データベース121は音素データを保存する。音声合成目的デ
15 ータ保存部122は音声合成目的データを保存する。

このように構成された携帯端末装置について、その動作を図3を用いて説明する。図3は図2の端末装置201の動作を示すフローチャートである。

ユーザが操作部106にて装置201の電源をオンすると、システム制御部101は記憶装置インタフェース部103に指令を出し、記憶装置202が端末装
20 置201に接続されているか否かを判定する(S1)。接続されていない場合は、記憶部105からフォントデータを取り出し、表示部107に「メモリカードを挿入してください」等のようなメッセージを表示し、ユーザに対して端末装置201に記憶装置202の接続を行うように促す(S2)。接続されている場合は、「操作メニュー1、再生2、プログラムアップデート3、読み上げ対象データダ
25 ウンロード4、キャラクタ音声データダウンロード・・・」等のメッセージを表示部107に表示する(S3)。ユーザがプログラムアップデートを選択すると(S4)、システム制御部101は、通信処理部110を通じて、端末装置201が公衆回線と接続されているか否かを判定し(S6)、接続されていない場合

は、「回線と接続してください」等の接続を促すメッセージを表示部107に表示する(S7)。接続されている場合は、通信処理部110を通じ、インターネット上のサーバ装置205(図1参照)と接続する(S8)。

接続が完了したら、端末装置201のシステム制御部101は、サーバ装置205に最新の音声合成処理プログラムのダウンロードを要求し、サーバ装置205から音声合成処理プログラムが送られてくると、音声合成処理プログラムを記憶部105に記憶して更新し(S9)、ダウンロード完了のメッセージを表示して、サーバ装置205との呼を切断する(S11)。

読み上げ対象データ(音声合成目的データ)やキャラクタ音声データ(音素データ)のダウンロードが選択された場合(S5)にサーバ装置205に接続してデータをダウンロードするまでの手順は、上述のプログラムアップデートと同じ手順であるが、読み上げ対象データやキャラクタ音声データのダウンロード時において、システム制御部101は、ダウンロードしたデータを、記憶装置インタフェース部103、記憶装置202内の端末装置インタフェース部120を通じて、記憶装置202に記憶する(S10)。そして、ダウンロード完了のメッセージを表示し、サーバ装置205との呼を切断する(S11)。

操作部106から再生が選択されると(S12)、システム制御部101は、表示部107に読み上げ対象データとキャラクタ音声のリストを表示し(S13)、ユーザに読み上げ対象データとキャラクタ音声データとを選択させる。ユーザが選択したら、記憶装置インタフェース部103に、記憶装置202内にある該当の音声合成目的データを読み出す指令を出し、記憶装置インタフェース部103は記憶装置202内にある端末装置インタフェース部120とやり取りをしながら音声合成目的データを読み出し、端末装置201内の記憶部105に記録する(S14)。次に、システム制御部101は、音声合成処理部102に指令を出し、処理を開始させる。音声合成処理部102は、記憶部105から音声合成目的データを順次読み出しながら解析を行い、データに最も適する音素データを記憶装置202から読み出して繋ぎ合わせ、合成音声データを作成し(S15)、その合成音声データを音声出力処理部104が処理できるデータに変換し

て音声出力処理部104に渡す。音声合成処理部102は、ユーザが操作部106で停止のボタンを押して、システム制御部101から処理中止の指示があるまで、この処理を繰り返す。音声出力処理部104は、音声合成処理部102から受け取ったデータをフォーマット変換し、スピーカまたはヘッドフォン108に出力する(S16)。

以上のように本実施の形態によれば、端末装置201は、全体を制御するシステム制御部101と、音声合成目的データと音素データとから音声合成処理プログラムに基づいて音声合成を行う音声合成処理部102と、音声合成処理プログラムを記憶する記憶部105と、記憶装置202との間でデータの送受信を行うための記憶装置インタフェース部103と、公衆回線に接続するための通信処理部110とを有し、記憶装置202は、音素データを格納する音素データベース121と、音声合成目的データを格納する音声合成目的データ122と、端末装置との間でデータの送受信を行うための端末装置インタフェース部120とを有することにより、ユーザは所望するキャラクタ音声にてテキストデータ等の朗読を聴くことができる。

(実施の形態2)

図4は本実施の形態の携帯端末装置を有する通信システムを示す構成図である。図4において、端末装置401は表示部や操作部、ヘッドフォン・スピーカ等の音声出力部などを備える。メモ리카ード等の記憶装置402はテキストデータ等の音声合成目的データ403を記憶する。記憶装置402は端末装置401とは脱着可能である。インターネット上のサーバ装置405は音声合成目的データ403や音素データ404、音声合成処理プログラム406を提供する。

図4においては、キャラクタ音声の音素データは端末装置401内の記憶部に搭載されている。ユーザは、インターネット上にあるサーバ装置405から自分の所望するキャラクタ音声の音素データを端末装置401内の記憶部に記憶させることもできる。また、最新の音声合成処理プログラムをダウンロードして、記憶部にあるプログラムを更新することも可能である。ユーザは、読み上げ対象の

音声合成目的データが格納された記憶装置402を端末装置401に挿入して操作することで、自分の所望のキャラクタ音声にて合成音声を聴くことができる。

図5は、図4の端末装置401と記憶装置402とを詳細に示すブロック図である。図5において、システム制御部501～通信処理部510、端末装置インタフェース部520、音声合成目的データ521は、図2のシステム制御部101～通信処理部110、端末装置インタフェース部120、音声合成目的データ122と同様のものである、その説明は省略する。

このように構成された携帯端末装置について、その動作を図6を用いて説明する。図6は図5の端末装置401の動作を示すフローチャートである。

ユーザが操作部506にて装置401の電源をオンすると、システム制御部501は記憶装置インタフェース部503に指令を出し、記憶装置402が端末装置401に接続されているか否かを判定する(S21)。接続されていない場合は、記憶部505からフォントデータを取り出し、表示部507に「メモリカードを挿入してください」等のようなメッセージを表示し、ユーザに対して端末装置401に記憶装置402の接続を行うように促す(S22)。接続されている場合は、「操作メニュー1、再生2、プログラムアップデート3、読み上げ対象データダウンロード4、キャラクタ音声データダウンロード・・・」等のメッセージを表示部507に表示する(S23)。ユーザがプログラムアップデートまたはキャラクタ音声データダウンロードを選択すると(S24)、システム制御部501は、通信処理部510を通じて、端末装置401が公衆回線と接続されているか否かを判定し(S26)、接続されていない場合は、「回線と接続してください」等の接続を促すメッセージを表示部507に表示する(S27)。接続されている場合は、通信処理部510を通じ、インターネット上のサーバ装置405(図4参照)と接続する(S28)。

接続が完了したら、端末装置401のシステム制御部501は、サーバ装置405に最新の音声合成処理プログラムまたはキャラクタ音声データ(該当する音楽データ)のダウンロードを要求し、サーバ装置405から音声合成処理プログラムが送られてくると、音声合成処理プログラムを記憶部505に記憶し、音声

合成処理プログラムを更新する（Ｓ２９）。音素データが送られてきた場合も同様な手順で行われ、記憶部５０５に記憶される（Ｓ２９）。そして、ダウンロード完了のメッセージを表示して、サーバ装置４０５との呼を切断する（Ｓ３１）。

読み上げ対象データ（音声合成目的データ）のダウンロードが選択された場合（Ｓ２５）にサーバ装置４０５に接続してデータをダウンロードするまでの手順は、上述のプログラムアップデートやキャラクタ音声ダウンロードと同じ手順であるが、読み上げ対象データのダウンロード時において、システム制御部５０１は、ダウンロードしたデータを、記憶装置インタフェース部５０３、記憶装置４０２内の端末装置インタフェース部５２０を通じて、記憶装置４０２に記憶する（Ｓ３０）。そして、ダウンロード完了のメッセージを表示し、サーバ装置４０５との呼を切断する（Ｓ３１）。

操作部５０６から再生が選択されると（Ｓ３２）、システム制御部５０１は、表示部５０７に読み上げ対象データとキャラクタ音声データのリストを表示し（Ｓ３３）、ユーザに読み上げ対象データとキャラクタ音声データとを選択させる。ユーザが選択したら、記憶装置インタフェース部５０３に、記憶装置４０２内にあるユーザが選択した音声合成目的データを読み出す指令を出す。記憶装置インタフェース部５０３は、記憶装置４０２内にある端末装置インタフェース部５２０とやり取りをしながら音声合成目的データを読み出し、端末装置４０１内の記憶部５０５に記録する（Ｓ３４）。

次に、システム制御部５０１は、音声合成処理部５０２に指令を出し、処理を開始させる。音声合成処理部５０２は、記憶部５０５から音声合成目的データを順次読み出しながら解析を行い、各データに最も適する音素データを記憶部５０５から読み出して繋ぎ合わせ、合成音声データを作成し（Ｓ３５）、その合成音声データを音声出力処理部５０４が処理できるデータに変換して音声出力処理部５０４に渡す。音声合成処理部５０２は、ユーザが操作部５０６で停止のボタンを押して、システム制御部５０１から処理中止の指示があるまで、この処理を繰り返す。音声出力処理部５０４は、音声合成処理部５０２から受け取ったデータをフォーマット変換し、スピーカまたはヘッドフォン５０８に出力する（Ｓ３

6)。

以上のように本実施の形態によれば、端末装置401は、全体を制御するシステム制御部501と、音声合成目的データと音素データとから音声合成処理プログラムに基づいて音声合成を行う音声合成処理部502と、音声合成処理プログラムと音素データとを記憶する記憶部505と、記憶装置402との間でデータの送受信を行うための記憶装置インタフェース部503と、公衆回線に接続するための通信処理部510とを有し、記憶装置402は、音声合成目的データを格納する音声合成目的データ521と、端末装置との間でデータの送受信を行うための端末装置インタフェース部520とを有することにより、ユーザは所望するキャラクタ音声にてテキストデータ等の朗読を聴くことができる。

(実施の形態3)

図7は本実施の形態の携帯端末装置を有する通信システムを示す構成図である。図7において、端末装置701は表示部や操作部、ヘッドフォン・スピーカ等の音声出力部などを備える。メモ리카ード等の記憶装置702はキャラクタ音声の音素データ704を記憶する。記憶装置702は端末装置701とは脱着可能である。インターネット上のサーバ装置である705は読み上げ対象の音声合成目的データ703や音素データ704、音声合成処理プログラム706を提供する。

図7においては、読み上げ対象データである音声合成目的データは端末装置701内の記憶部に搭載されている。ユーザは、インターネット上にあるサーバ装置705から自分の所望する音声合成目的データを端末装置701内の記憶部に記憶させることもできる。また、最新の音声合成処理プログラムをダウンロードして記憶部にある音声合成処理プログラムの更新も可能である。ユーザは、キャラクタ音声の音素データが格納された記憶装置702を端末装置701に挿入して操作することで、自分の所望のキャラクタ音声にて合成音を聴くことができる。

図8は、図7の端末装置701と記憶装置702とを詳細に示すブロック図である。図8において、システム制御部801～通信処理部810、端末装置イン

タフェース部 820、音素データベース 821 は、図 2 のシステム制御部 101
～通信処理部 110、端末装置インタフェース部 120、音素データベース 12
1 と同様のものであるため、その説明は省略する。このように構成された携帯端
末装置について、その動作を図 9 を用いて説明する。図 9 は図 8 の端末装置 70
5 1 の動作を示すフローチャートである。

ユーザが操作部 806 にて装置 701 の電源をオンすると、システム制御部 8
01 は記憶装置インタフェース部 803 に指令を出し、記憶装置 702 が端末装
置 701 に接続されているか否かを判定する (S41)。接続されていない場合
は、記憶部 805 からフォントデータを取り出し、表示部 807 に「メモリカー
ドを挿入してください」等のようなメッセージを表示し、ユーザに対して端末装
置 701 に記憶装置 702 の接続を行うように促す (S42)。接続されている
場合は、「操作メニュー 1、再生 2、プログラムアップデート 3、読み上げ対象
データダウンロード 4、キャラクタ音声データダウンロード・・・」等のメッセ
ージを表示部 807 に表示する (S43)。ユーザがプログラムアップデートまた
は読み上げ対象データダウンロードを選択すると (S44)、システム制御部 8
01 は、通信処理部 810 を通じて、端末装置 701 が公衆回線と接続されてい
るか否かを判定し (S46)、接続されていない場合は、「回線と接続してくださ
い」等の接続を促すメッセージを表示部 807 に表示する (S47)。接続され
ている場合は、通信処理部 810 を通じ、インターネット上のサーバ装置 705
(図 7 参照) と接続する (S38)。

接続が完了すると、端末装置 701 のシステム制御部 801 は、サーバ装置 7
05 に最新の音声合成処理プログラムまたは読み上げ対象データ (該当する音声
合成目的データ) のダウンロードを要求し、サーバ装置 705 から音声合成処理
プログラムが送られてくると、音声合成処理プログラムを記憶部 805 に記憶し、
音声合成処理プログラムを更新する (S49)。音声合成目的データが送られて
きた場合も同様な手順で行われ、記憶部 805 に記憶される (S49)。そして、
ダウンロード完了のメッセージを表示して、サーバ装置 705 との呼を切断する
(S51)。

キャラクタ音声データ（該当する音素データ）のダウンロードが選択された場合（S45）にサーバ装置705に接続してデータをダウンロードするまでの手順は、上述のプログラムアップデートや読み上げ対象データダウンロードと同じ手順であるが、キャラクタ音声データのダウンロード時において、システム制御部801は、ダウンロードしたデータを、記憶装置インタフェース部803、記憶装置702内の端末装置インタフェース部820を通じて、記憶装置702に記憶する（S50）。そして、ダウンロード完了のメッセージを表示し、サーバ装置705との呼を切断する（S51）。

操作部806から再生が選択されると（S52）、システム制御部801は、表示部807に読み上げ対象データとキャラクタ音声データのリストを表示し（S53）、ユーザに読み上げ対象データとキャラクタ音声データとを選択させる。ユーザが選択したら、システム制御部801は、音声合成処理部802に指令を出し、処理を開始させる。音声合成処理部802は、記憶部805から音声合成目的データを順次読み出しながら解析を行い、各データに最も適する音素データを記憶装置702から読み出して繋ぎ合わせ、合成音声データを作成し（S54）、その合成音声データを音声出力処理部804が処理できるデータに変換して音声出力処理部804に渡す。音声合成処理部802は、ユーザが操作部806で停止のボタンを押して、システム制御部801から処理中止の指示があるまで、この処理を繰り返す。音声出力処理部804は、音声合成処理部802から受け取ったデータをフォーマット変換し、スピーカまたはヘッドホン508に出力する（S55）。

以上のように本実施の形態によれば、端末装置701は、全体を制御するシステム制御部801と、音声合成目的データと音素データとから音声合成処理プログラムに基づいて音声合成を行う音声合成処理部802と、音声合成処理プログラムと音声合成目的データとを記憶する記憶部805と、記憶装置702との間でデータの送受信を行うための記憶装置インタフェース部803と、公衆回線に接続するための通信処理部810とを有し、記憶装置702は、音素データを格納する音素データベース821と、端末装置701との間でデータの送受信を行

うための端末装置インタフェース部820とを有することにより、ユーザは所望するキャラクタ音声にてテキストデータ等の朗読を聴くことができる。

(実施の形態4)

- 5 図10は本実施の形態の携帯端末装置を有する通信システムを示す構成図である。図10において、端末装置1001は表示部や操作部、ヘッドフォン・スピーカ等の音声出力部などを備える。メモリカード等の記憶装置1002は読み上げ対象の音声合成目的データ1003とキャラクタ音声の音素データ1004と音声合成処理プログラム1006とを記憶する。記憶装置1002は端末装置1001とは脱着可能である。インターネット上のサーバ装置1005は読み上げ対象の音声合成目的データ1003や音素データ1004、音声合成処理プログラム1006を提供する。

- 10 図10において、ユーザは、音声合成処理プログラムと読み上げ対象の音声合成目的データとキャラクタ音声の音素データとを格納する記憶装置1002を端末装置1001に挿入して操作することで、自分の所望のキャラクタ音声にてテキストの朗読を聴くことができる。

- 15 図11は、図10の端末装置1001と記憶装置1002とを詳細に示すブロック図である。図11において、システム制御部1101、記憶装置インタフェース部1103、音声出力処理部1104～通信処理部1110、端末装置インタフェース部1120、音素データベース1122、音声合成目的データ1123は、図2のシステム制御部101、記憶装置インタフェース部103、音声出力処理部104～通信処理部110、端末装置インタフェース部120、音素データベース121、音声合成目的データ122と同様のものである。その説明は省略する。1121は音声合成処理プログラムを保持する音声合成処理部である。

20 このように構成された携帯端末装置について、その動作を図12を用いて説明する。図12は端末装置1001の動作を示すフローチャートである。

ユーザが操作部1106にて装置1001の電源をオンすると、システム制御

部1101は記憶装置インタフェース部1103に指令を出し、記憶装置1002が端末装置1001に接続されているか否かを判定する(S61)。接続されていない場合は、記憶部1105からフロントデータを取り出し、表示部1107に「メモリカードを挿入してください」等のようなメッセージを表示し、ユーザに対して端末装置1001に記憶装置1002の接続を行うように促す(S62)。接続されている場合は、「操作メニュー1、再生2、プログラムアップデート3、読み上げ対象データダウンロード4、キャラクタ音声データダウンロード・・・」等のメッセージを表示部1107に表示する(S63)。ユーザが再生以外を選択すると(S65)、システム制御部1101は、通信処理部1110を通じて、端末装置1001が公衆回線と接続されているか否かを判定し(S66)、接続されていない場合は、「回線と接続してください」等の接続を促すメッセージを表示部1107に表示する(S67)。接続されている場合は、通信処理部1110を通じ、インターネット上のサーバ装置1005(図10参照)と接続する(S68)。

接続が完了すると、端末装置1001のシステム制御部1101は、選択された操作内容に従い、サーバ装置1005に最新の音声合成処理プログラム、キャラクタ音声データ(該当する音素データ)または読み上げ対象データ(音声合成目的データ)のダウンロードを要求する。そして、サーバ装置1005から音声合成処理プログラムが送られてくると、音声合成処理プログラムを記憶部1105に記憶し、音声合成処理プログラムを更新する(S69)。音素データや読み上げ対象データが送られてきた場合も同様な手順で行われ、記憶部1105に記憶される(S70)。そして、ダウンロード完了のメッセージを表示して、サーバ装置1005との呼を切断する(S71)。

操作部1106から再生が選択されると(S72)、システム制御部1101は、表示部1107に読み上げ対象データとキャラクタ音声データのリストを表示し(S73)、ユーザに読み上げ対象データとキャラクタ音声データとを選択させる。ユーザが選択したら、記憶装置インタフェース部1103に、記憶装置1002内にある音声合成処理プログラムとユーザが選択した音声合成目的デー

タとを読み出す指令を出す。記憶装置インタフェース部1103は、記憶装置1002内にある端末装置インタフェース部1120とやり取りをしながら音声合成処理プログラムと音声合成目的データとを読み出し、端末装置1001内の記憶部1105に記録する(S74)。

- 5 次に、システム制御部1101は、音声合成処理部1102に指令を出し、処理を開始させる。音声合成処理部1102は、記憶部1105から音声合成目的データを順次読み出しながら解析を行い、各データに最も適する音素データを記憶装置1002から読み出して繋ぎ合わせ、合成音声データを作成し(S75)、その合成音声データを音声出力処理部1104が処理できるデータに変換して音声出力処理部1104に渡す。音声合成処理部1102は、ユーザが操作部1106で停止のボタンを押して、システム制御部1101から処理中止の指示があるまで、この処理を繰り返す。音声出力処理部1104は、音声合成処理部1102から受け取ったデータをフォーマット変換し、スピーカまたはヘッドフォン1108に出力する(S76)。
- 10 15 以上のように本実施の形態によれば、端末装置1001は、全体を制御するシステム制御部1101と、音声合成処理プログラムを記憶する記憶部1105と、記憶装置1002との間でデータの送受信を行うための記憶装置インタフェース部1103と、公衆回線に接続するための通信処理部1110とを有し、記憶装置1002は、音声合成目的データと音素データとから音声合成処理プログラム
- 20 25 に基づいて音声合成を行うと共に音声合成処理プログラムを格納する音声合成処理部1121と、音素データを格納する音素データベース1122と、音声合成目的データを格納する音声合成目的データ1123と、端末装置1001との間でデータの送受信を行うための端末装置インタフェース部1120とを有することにより、ユーザは所望するキャラクタ音声にてテキストデータ等の朗読を聴くことができる。

(実施の形態5)

図13は本実施の形態の携帯端末装置を有する通信システムを示す構成図であ

る。図13において、端末装置1301は表示部や操作部、ヘッドフォン・スピーカ等の音声出力部などを備える。メモリカード等の記憶装置1302は読み上げ対象の音声合成目的データ1303と音声合成処理プログラム1306とを記憶する。記憶装置1302は端末装置1301とは脱着可能である。インターネット上のサーバ装置1305は読み上げ対象の音声合成目的データ1303や音素データ1304、音声合成処理プログラム1306を提供する。

図13においては、キャラクタ音声の音素データは端末装置1301内の記憶部に搭載されている。ユーザは、インターネット上にあるサーバ装置1305から自分の所望するキャラクタ音声の音素データを端末装置1301内の記憶部に記憶させることもできる。また、最新の音声合成処理プログラム1306をダウンロードして、記憶装置1302にある音声合成処理プログラムの更新も可能である。ユーザは、読み上げ対象の音声合成目的データが格納された記憶装置1302を端末装置1301に挿入して操作することで、自分の所望のキャラクタ音声にて合成音声を聴くことができる。

図14は、図13の端末装置1301と記憶装置1302とを詳細に示すブロック図である。図14において、システム制御部1401、記憶装置インタフェース部1403、音声出力処理部1404～通信処理部1410、端末装置インタフェース部1420、音声合成目的データ1423は、図2のシステム制御部101、記憶装置インタフェース部103、音声出力処理部104～通信処理部110、端末装置インタフェース部120、音声合成目的データ122と同様のものである。その説明は省略する。1421は音声合成処理プログラムを保持する音声合成処理部である。

このように構成された携帯端末装置について、その動作を図15を用いて説明する。図15は端末装置1301の動作を示すフローチャートである。

ユーザが操作部1406にて装置1301の電源をオンすると、システム制御部1401は記憶装置インタフェース部1403に指令を出し、記憶装置1302が端末装置1301に接続されているか否かを判定する(S81)。接続されていない場合は、記憶部1405からフォントデータを取り出し、表示部140

7に「メモリカードを挿入してください」等のようなメッセージを表示し、ユーザに対して端末装置1301に記憶装置1302の接続を行うように促す(S82)。接続されている場合は、「操作メニュー1、再生2、プログラムアップデート3、読み上げ対象データダウンロード4、キャラクタ音声データダウンロード・・」等のメッセージを表示部1407に表示する(S83)。ユーザがキャラクタ音声データダウンロードを選択すると(S84)、システム制御部1401は、通信処理部1410を通じて、端末装置1301が公衆回線と接続されているか否かを判定し(S86)、接続されていない場合は、「回線と接続してください」等の接続を促すメッセージを表示部1407に表示する(S87)。接続されている場合は、通信処理部1410を通じ、インターネット上のサーバ装置1305(図13参照)と接続する(S88)。

接続が完了すると、端末装置1301のシステム制御部1401は、サーバ装置1305にキャラクタ音声データ(該当する音素データ)のダウンロードを要求し、サーバ装置1305からキャラクタ音声データが送られてくると、キャラクタ音声データを記憶部1405に記憶する(S89)。そして、ダウンロード完了のメッセージを表示して、サーバ装置1305との呼を切断する(S91)。

プログラムアップデートまたは読み上げ対象データ(音声合成目的データ)のダウンロードが選択された場合(S85)にサーバ装置1305に接続してデータをダウンロードするまでの手順は、上述のキャラクタ音声ダウンロードと同じ手順であるが、この場合、システム制御部1401は、ダウンロードしたデータを、記憶装置インタフェース部1403、記憶装置1302内の端末装置インタフェース部1420を通じて、記憶装置1302に記憶する(S90)。そして、ダウンロード完了のメッセージを表示し、サーバ装置1305との呼を切断する(S91)。

操作部1406から再生が選択されると(S92)、システム制御部1401は、表示部1407に読み上げ対象データとキャラクタ音声データのリストを表示し(S93)、ユーザに読み上げ対象データとキャラクタ音声データとを選択させる。ユーザが選択したら、記憶装置インタフェース部1403に、記憶装置

1302内にあるユーザが選択した音声合成目的データを読み出す指令を出す。

記憶装置インタフェース部1403は、記憶装置1302内にある端末装置インタフェース部1420とやり取りをしながら音声合成目的データを読み出し、端末装置1301内の記憶部1405に記録する(S94)。

次に、システム制御部1401は、音声合成処理部1402に指令を出し、処理を開始させる。音声合成処理部1402は、記憶部1405から音声合成目的データを順次読み出し、ながら解析を行い、各データに最も適する音素データを記憶部1405から読み出して繋ぎ合わせ、合成音声データを作成し(S95)、その合成音声データを音声出力処理部1404が処理できるデータに変換して音声出力処理部1404に渡す。

音声合成処理部1402は、ユーザが操作部1406で停止のボタンを押して、システム制御部1401から処理中止の指示があるまで、この処理を繰り返す。音声出力処理部1404は、音声合成処理部1402から受け取ったデータをフォーマット変換し、スピーカまたはヘッドフォン1408に出力する(S96)。

以上のように本実施の形態によれば、端末装置1301は、全体を制御するシステム制御部1401と、音声合成処理プログラムと音素データとを記憶する記憶部1405と、記憶装置1302との間でデータの送受信を行うための記憶装置インタフェース部1403と、公衆回線に接続するための通信処理部1410とを有し、記憶装置1302は、音声合成目的データと音素データとから音声合成処理プログラムに基づいて音声合成を行うと共に音声合成処理プログラムを格納する音声合成処理部1421と、音声合成目的データを格納する音声合成目的データ1423と、端末装置との間でデータの送受信を行うための端末装置インタフェース部1420とを有することにより、ユーザは所望するキャラクタ音声にてテキストデータ等の朗読を聴くことができる。

(実施の形態6)

図16は本実施の形態の携帯端末装置を有する通信システムを示す構成図である。図16において、端末装置1601は表示部や操作部、ヘッドフォン・スピー

一カ等の音声出力部などを備える。メモリカード等の記憶装置1602は音素データ1604と音声合成処理プログラム1606とを記憶する。記憶装置1602は端末装置1601とは脱着可能である。インターネット上のサーバ装置1605は読み上げ対象の音声合成目的データ1603や音素データ1604、音声合成処理プログラム1606を提供する。

図16においては、読み上げ対象データの音声合成目的データは端末装置1601内の記憶部に搭載されている。ユーザは、インターネット上にあるサーバ装置1605から自分の所望する音声合成目的データを端末装置1601内の記憶部に記憶させることもできる。また、最新の音声合成処理プログラムをダウンロードして記憶装置にある音声合成処理プログラムの更新も可能である。ユーザは、音声キャラクタの音素データが格納された記憶装置1602を端末装置1601に挿入して操作することで、自分の所望のキャラクタ音声にて合成音を聴くことができる。

図17は、図16の端末装置1601と記憶装置1602とを詳細に示すブロック図である。図17において、システム制御部1701、記憶装置インタフェース部1703、音声出力処理部1704～通信処理部1710、端末装置インタフェース部1720、音素データベース1722は、図2のシステム制御部101、記憶装置インタフェース部103、音声出力処理部104～通信処理部110、端末装置インタフェース部120、音素データベース121と同様のものである。1721は音声合成処理プログラムを保持する音声合成処理部である。

このように構成された携帯端末装置について、その動作を図18を用いて説明する。図18は端末装置1601の動作を示すフローチャートである。ユーザが操作部1706にて装置1601の電源をオンすると、システム制御部1701は記憶装置インタフェース部1703に指令を出し、記憶装置1602が端末装置1601に接続されているか否かを判定する(S101)。接続されていない場合は、記憶部1705からフォントデータを取り出し、表示部1707に「メモリカードを挿入してください」等のようなメッセージを表示し、ユーザに対し

て端末装置1601に記憶装置1602の接続を行うように促す(S102)。接続されている場合は、「操作メニュー1、再生2、プログラムアップデート3、読み上げ対象データダウンロード4、キャラクタ音声データダウンロード・・・」等のメッセージを表示部1707に表示する(S103)。

- 5 ユーザが読み上げ対象データダウンロードを選択すると(S104)、システム制御部1701は、通信処理部1710を通じて、端末装置1601が公衆回線と接続されているか否かを判定し(S106)、接続されていない場合は、「回線と接続してください」等の接続を促すメッセージを表示部1707に表示する(S107)。接続されている場合は、通信処理部1710を通じ、インターネット上のサーバ装置1605(図16参照)と接続する(S108)。

- 10 接続が完了すると、端末装置1601のシステム制御部1701は、サーバ装置1605に読み上げ対象データ(該当する音声合成目的データ)のダウンロードを要求し、サーバ装置1605から読み上げ対象データが送られてくると、読み上げ対象データを記憶部1705に記憶する(S109)。そして、ダウンロード完了のメッセージを表示して、サーバ装置1605との呼を切断する(S111)。

- 15 プログラムアップデートまたはキャラクタ音声データ(該当する音素データ)のダウンロードが選択された場合(S105)にサーバ装置1605に接続してデータをダウンロードするまでの手順は、上述の読み上げ対象データダウンロードと同じ手順であるが、この場合、システム制御部1701は、ダウンロードしたデータを、記憶装置インタフェース部1703、記憶装置1602内の端末装置インタフェース部1720を通じて、記憶装置1602に記憶する(S110)。そして、ダウンロード完了のメッセージを表示し、サーバ装置1605との呼を切断する(S111)。

- 25 操作部1706から再生が選択されると(S112)、システム制御部1701は、表示部1707に読み上げ対象データとキャラクタ音声データのリストを表示し(S113)、ユーザに読み上げ対象データとキャラクタ音声データとを選択させる。ユーザが選択したら、システム制御部1701は、音声合成処理部

1721に指令を出し、処理を開始させる。音声合成処理部1702は、記憶部1705から音声合成目的データを順次読み出しながら解析を行い、各データに最も適する音素データを記憶装置1602から読み出して繋ぎ合わせ、合成音声データを作成し（S114）、その合成音声データを音声出力処理部1704が

5 処理できるデータに変換して音声出力処理部1704に渡す。音声合成処理部1702は、ユーザが操作部1706で停止のボタンを押して、システム制御部1701から処理中止の指示があるまで、この処理を繰り返す。音声出力処理部1704は、音声合成処理部1702から受け取ったデータをフォーマット変換し、スピーカまたはヘッドホン1708に出力する（S115）。

10 以上のように本実施の形態によれば、端末装置1601は、全体を制御するシステム制御部1701と、音声合成処理プログラムと音声合成目的データとを記憶する記憶部1705と、記憶装置1602との間でデータの送受信を行うための記憶装置インタフェース部1703と、公衆回線に接続するための通信処理部1710とを有し、記憶装置1602は、音声合成目的データと音素データとから

15 音声合成処理プログラムに基づいて音声合成を行うと共に音声合成処理プログラムを格納する音声合成処理部1721と、音素データを格納する音素データベース1722と、端末装置1601との間でデータの送受信を行うための端末装置インタフェース部1720とを有することにより、ユーザは所望するキャラクター音声にてテキストデータ等の朗読を聴くことができる。

（実施の形態7）

図19は本実施の形態の携帯端末装置を有する通信システムを示す構成図である。図19において、端末装置1901は表示部や操作部、ヘッドホン・スピーカ等の音声出力部などを備える。メモリカード等の記憶装置1902は音声合

25 成処理プログラム1906を記憶する。記憶装置1902は端末装置1901とは脱着可能である。インターネット上のサーバ装置1905は読み上げ対象の音声合成目的データ1903や音素データ1904、音声合成処理プログラム1906を提供する。

図19においては、読み上げ対象データの音声合成目的データおよびキャラクタ音声の音素データは端末装置1901内の記憶部に搭載されている。ユーザは、インターネット上にあるサーバ装置1905から自分の所望する音声合成目的データや音素データをダウンロードして端末装置1901内の記憶部に記憶させることもできる。また、最新の音声合成処理プログラムをダウンロードして記憶装置1902にある音声合成処理プログラムの更新も可能である。ユーザは、音声合成処理プログラムが搭載された記憶装置1902を端末装置1901に挿入して操作することで、自分の所望のキャラクタ音声にて合成音を聴くことができる。

図20は、図19の端末装置1901と記憶装置1902とを詳細に示すブロック図である。図20において、システム制御部2001、記憶装置インタフェース部2003、音声出力処理部2004～通信処理部2010、端末装置インタフェース部2020は、図2のシステム制御部101、記憶装置インタフェース部103、音声出力処理部104～通信処理部110、端末装置インタフェース部120と同様のものである。2021は音声合成処理プログラムを保持する音声合成処理部である。

このように構成された携帯端末装置について、その動作を図21を用いて説明する。図21は端末装置1901の動作を示すフローチャートである。ユーザが操作部2006にて装置1901の電源をオンすると、システム制御部2001は記憶装置インタフェース部2003に指令を出し、記憶装置1902が端末装置1901に接続されているか否かを判定する(S121)。接続されていない場合は、記憶部2005からフォントデータを取り出し、表示部2007に「メモリカードを挿入してください」等のようなメッセージを表示し、ユーザに対して端末装置1901に記憶装置1902の接続を行うように促す(S122)。

接続されている場合は、「操作メニュー1、再生2、プログラムアップデート3、読み上げ対象データダウンロード4、キャラクタ音声データダウンロード・・・」等のメッセージを表示部2007に表示する(S123)。ユーザが読み上げ対象データやキャラクタ音声データのダウンロードを選択すると(S12

- 4)、システム制御部2001は、通信処理部2010を通じて、端末装置1901が公衆回線と接続されているか否かを判定し(S126)、接続されていない場合は、「回線と接続してください」等の接続を促すメッセージを表示部2007に表示する(S127)。接続されている場合は、通信処理部2010を通じて、インターネット上のサーバ装置1905(図19参照)と接続する(S128)。

- 接続が完了すると、端末装置1901のシステム制御部2001は、操作の選択に従い、サーバ装置1905に読み上げ対象データ(該当する音声合成目的データ)やキャラクタ音声データ(該当する音素データ)のダウンロードを要求し、
- 10 サーバ装置1905から読み上げ対象データやキャラクタ音声データが送られてくると、読み上げ対象データやキャラクタ音声データを記憶部2005に記憶する(S129)。そして、ダウンロード完了のメッセージを表示して、サーバ装置1905との呼を切断する(S131)。

- プログラムアップデートが選択された場合(S125)にサーバ装置1905
- 15 に接続してデータをダウンロードするまでの手順は、上述の読み上げ対象データやキャラクタ音声データのダウンロードと同じ手順であるが、この場合、システム制御部2001は、ダウンロードしたデータを、記憶装置インタフェース部2003、記憶装置1902内の端末装置インタフェース部2020を通じて、記憶装置1902に記憶し、音声合成処理プログラムを更新する(S130)。そ
- 20 して、ダウンロード完了のメッセージを表示し、サーバ装置1905との呼を切断する(S131)。

- 操作部2006から再生が選択されると(S132)、システム制御部2001は、表示部2007に読み上げ対象データとキャラクタ音声データのリストを表示し(S133)、ユーザに読み上げ対象データとキャラクタ音声データとを
- 25 選択させる。ユーザが選択したら、システム制御部2001は、記憶装置1902から音声合成処理プログラムを読み出し、記憶部2005に記憶する(S134)。次に、システム制御部2001は、音声合成処理部2021に指令を出し、処理を開始させる。音声合成処理部2021は、記憶部2005から音声合成目

的データを順次読み出しながら解析を行い、各データに最も適する音素データを記憶部2005から読み出して繋ぎ合わせ、合成音声データを作成し（S135）、その合成音声データを音声出力処理部2004が処理できるデータに変換して音声出力処理部2004に渡す。音声合成処理部2021は、ユーザが操作部2006で停止のボタンを押して、システム制御部2001から処理中止の指示があるまで、この処理を繰り返す。音声出力処理部2004は、音声合成処理部2021から受け取ったデータをフォーマット変換し、スピーカまたはヘッドフォン2008に出力する（S136）。

以上のように本実施の形態によれば、端末装置1901は、全体を制御するシステム制御部2001と、音声合成処理プログラムと音声合成目的データと音素データとを記憶する記憶部2005と、記憶装置1902との間でデータの送受信を行うための記憶装置インタフェース部2003と、公衆回線に接続するための通信処理部2010とを有し、記憶装置1902は、音声合成目的データと音素データとから音声合成処理プログラムに基づいて音声合成を行うと共に音声合成処理プログラムを格納する音声合成処理部2021と、端末装置1901との間でデータの送受信を行うための端末装置インタフェース部2020とを有することにより、ユーザは所望するキャラクタ音声にてテキスト

（実施の形態8）

図22は本実施の形態の携帯端末装置を有する通信システムを示す構成図である。図22において、端末装置2201は表示部や操作部、ヘッドホン・スピーカ等の音声出力部などを備える。メモリカード等の記憶装置2202は端末装置2201とは脱着可能である。インターネット上のサーバ装置2205は読み上げ対象の音声合成目的データ2203や音素データ2204、音声合成処理プログラム2206を提供する。

図22においては、読み上げ対象データの音声合成目的データおよびキャラクタ音声の音素データは端末装置1901内の記憶部に搭載されている。ユーザは、インターネット上にあるサーバ装置1905から自分の所望する音声合成目的デ

ータや音素データをダウンロードして端末装置1901内の記憶部に記憶させることもできる。また、最新の音声合成処理プログラムをダウンロードして端末装置1901にある音声合成処理プログラムの更新も可能である。ユーザは、端末装置を操作し、自分の所望のキャラクタ音声にて合成音声を聴くことができる。

- 5 図23は、図22の端末装置2201と記憶装置2202とを詳細に示すブロック図である。図23において、システム制御部2301、記憶装置インタフェース部2303、音声出力処理部2304～通信処理部2310、端末装置インタフェース部2320は、図2のシステム制御部101、記憶装置インタフェース部103、音声出力処理部104～通信処理部110、端末装置インタフェース部120と同様のものであるため、その説明は省略する

このように構成された携帯端末装置について、その動作を図24を用いて説明する。図24は端末装置2201の動作を示すフローチャートである。ユーザが操作部2306にて装置2201の電源をオンすると、システム制御部2301は「操作メニュー1、再生2、プログラムアップデート3、読み上げ対象データダウンロード4、キャラクタ音声データダウンロード・・・」等のメッセージを表示部2307に表示する（S2401）。ユーザが読み上げ対象データやキャラクタ音声データのダウンロードまたはプログラムアップデートを選択すると（S2402）、システム制御部2301は、通信処理部2310を通じて、端末装置2201が公衆回線と接続されているか否かを判定し（S2403）、接続されていない場合は、「回線と接続してください」等の接続を促すメッセージを表示部2307に表示する（S2404）。接続されている場合は、通信処理部2310を通じ、インターネット上のサーバ装置2205（図22参照）と接続する（S2405）。

接続が完了すると、端末装置2201のシステム制御部2301は、操作の選択に従い、サーバ装置2205に読み上げ対象データ（該当する音声合成目的データ）やキャラクタ音声データ（該当する音素データ）や最新の音声合成処理プログラムのダウンロードを要求し、サーバ装置2205から読み上げ対象データやキャラクタ音声データまたはプログラムが送られてくると、記憶部2305に

記憶する（Ｓ２４０６）。そして、ダウンロード完了のメッセージを表示して、サーバ装置２２０５との呼を切断する（Ｓ２４０７）。

操作部２３０６から再生が選択されると（Ｓ２４０８）、システム制御部２３０１は、表示部２３０７に読み上げ対象データとキャラクタ音声データのリストを表示し（Ｓ２４０９）、ユーザに読み上げ対象データとキャラクタ音声データとを選択させる。ユーザが選択したら、システム制御部２３０１は、音声合成処理部２３０２に指令を出し、処理を開始させる。音声合成処理部２３０２は、記憶部２３０５から音声合成目的データを順次読み出しながら解析を行い、各データに最も適する音素データを記憶部２３０５から読み出して繋ぎ合わせ、合成音声データを作成し（Ｓ２４１０）、その合成音声データを音声出力処理部２３０４が処理できるデータに変換して音声出力処理部２３０４に渡す。音声合成処理部２３０２は、ユーザが操作部２３０６で停止のボタンを押して、システム制御部２３０１から処理中止の指示があるまで、この処理を繰り返す。音声出力処理部２３０４は、音声合成処理部２３０２から受け取ったデータをフォーマット変換し、スピーカまたはヘッドフォン２３０８に出力する（Ｓ２４１１）。

以上のように本実施の形態によれば、端末装置２２０１は、全体を制御するシステム制御部２３０１と、音声合成処理プログラムと音声合成目的データと音素データとを記憶する記憶部２３０５と、記憶装置２２０２との間でデータの送受信を行うための記憶装置インタフェース部２２０３と、公衆回線に接続するための通信処理部２２１０と音声合成目的データと音素データとから音声合成処理プログラムに基づいて音声合成を行う音声合成処理部２３０２を有し、記憶装置２２０２は、端末装置２２０１との間でデータの送受信を行うための端末装置インタフェース部２３２０とを有することにより、ユーザは所望するキャラクタ音声にてテキストデータの朗読を聴くことができる。

（実施の形態９）

図２５は本実施の形態の携帯端末装置を有する通信システムを示す構成図である。図２５において、端末装置２５０１は表示部や操作部、ヘッドフォン・スピー

一カ等の音声出力部などを備える。メモリカード等の記憶装置2502はキャラクタ音声の音素データや読み上げ対象である音声合成目的データ、音声合成処理プログラムを記憶することが可能で、かつインターネット上の通信を実現する通信処理部を内蔵する。記憶装置2502は端末装置2501とは脱着可能である。

- 5 インターネット上のサーバ装置2505は読み上げ対象の音声合成目的データ2503や音素データ2504、音声合成処理プログラム2506を提供する。

- 10 図25においては、通信処理部が内蔵された記憶装置2502を端末装置2501に挿入して記憶装置2502の通信処理部を通じてインターネット上のサーバ装置2505から音声合成処理プログラムや音声合成目的データ、音素データを記憶装置2502上あるいは端末装置2501の記憶部に取り込んで操作することで、自分の所望のキャラクタ音声にて読み上げ対象の合成音を聴くことができる。

- 15 図26は、図25の端末装置2501と記憶装置2502とを詳細に示すブロック図である。図26において、システム制御部2601、記憶装置インタフェース部2603、音声合成処理部2602～電源部2609、端末装置インタフェース部2620は、図2のシステム制御部101、記憶装置インタフェース部103、音声出力処理部104～電源部109、端末装置インタフェース部120と同様のものであるので、その説明は省略する。記憶装置内蔵の通信処理部2610はインターネット上のサーバ装置との通信を実現する。

- 20 このように構成された携帯端末装置について、その動作を説明する。

- ユーザが操作部2606にて装置2501の電源をオンすると、システム制御部2601は記憶装置インタフェース部2603に指令を出し、記憶装置2502が端末装置2501に接続されているか否かを判定する。接続されていない場合は、記憶部2605からフォントデータを取り出し、表示部2607に「メモリカードを挿入してください」等のようなメッセージを表示し、ユーザに対して
25 端末装置2501に記憶装置2502の接続を行うように促す。接続されている場合は、「操作メニュー1、再生2、プログラムアップデート3、読み上げ対象データダウンロード4、キャラクタ音声データダウンロード・・・」等のメッセー

ジを表示部2607に表示する。ユーザが再生以外を選択すると、システム制御部2601は、通信処理部2610を通じて、端末装置2501が公衆回線と接続されているか否かを判定し、接続されていない場合は、「回線と接続してください」等の接続を促すメッセージを表示部2607に表示する。接続されている場合は、通信処理部2610を通じ、インターネット上のサーバ装置2505（図25参照）と接続する。

接続が完了すると、端末装置2501のシステム制御部2601は、操作の選択に従い、サーバ装置2505に読み上げ対象データ（該当する音声合成目的データ）やキャラクタ音声データ（該当する音素データ）、音声合成処理プログラムのダウンロードを要求し、サーバ装置2505から読み上げ対象データやキャラクタ音声データ、音声合成処理プログラムが送られてくると、各データを記憶部2605や記憶装置2502に記憶する。そして、ダウンロード完了のメッセージを表示して、サーバ装置2505との呼を切断する。

操作部2606から再生が選択された場合、音声合成音が出来上がるまでの内部の動作は、音声合成処理プログラムや音声合成目的データ、音素データの各々が記憶装置2502および端末装置2501の記憶部2605のどちらに記録されたかで異なってくるが、その記録形態は実施の形態1～実施の形態7（図3、図6、図9、図12、図15、図18、図21）のどれかに合致するので、その説明は省略する。

以上のように本実施の形態によれば、端末装置2501の通信処理部を無くして、記憶装置2502に通信処理部2610を備えたことにより、端末装置2501においては通信処理部を不要とすることができるので、端末装置2501を簡単な構成にすることができ、端末装置2501のコストダウンを図ることができる。

（実施の形態10）

図27は本実施の形態の携帯端末装置を有する通信システムを示す構成図である。図27において、端末装置2701は表示部や操作部、ヘッドフォン・スピー

一カ等の音声出力部などを備える。メモリカード等の記憶装置2702はキャラクタ音声の音素データや読み上げ対象である音声合成目的データ、音声合成処理プログラムを記憶することが可能で、かつインターネット上の通信を実現する通信処理部と通信処理部を制御する記憶制御部を内蔵する。記憶装置2702は端末装置2701とは脱着可能である。インターネット上のサーバ装置2705は読み上げ対象の音声合成目的データ2703や音素データ2704、音声合成処理プログラム2706を提供する。

図27の携帯端末装置2701、記憶装置2702においては、通信処理部と通信処理部を制御する記憶制御部が搭載され、記憶装置2702単体でインターネット上のサーバ装置2705から音声合成処理プログラムや音声合成目的データ、音素データがダウンロード可能であり、これらデータを記憶装置2702上あるいは端末装置2701の記憶部に取り込んで操作することで、自分の所望のキャラクタ音声にて読み上げ対象の合成音声を聴くことができる。

図28は、図27の端末装置2701と記憶装置2702とを詳細に示すブロック図である。図28において、システム制御部2801、記憶装置インタフェース部2803、音声合成処理部2802～電源部2809、端末装置インタフェース部2820は、図2のシステム制御部101、記憶装置インタフェース部103、音声出力処理部104～電源部109、端末装置インタフェース部120と同様のものである。その説明は省略する。記憶装置内蔵の通信処理部2810はインターネット上のサーバ装置2705との通信を実現する。記憶制御部2821は通信処理部2810を制御する。ユーザインタフェースを実現する記憶装置2702上にある記憶表示部2822は記憶装置単体で通信を行う際の情報を表示する。記憶操作部2823は記憶装置単体でのユーザインタフェースを実現する。

このように構成された携帯端末装置について、その動作を説明する。ユーザが記憶操作部2823にて記憶装置2702の電源をオンすると、記憶制御部2821は記憶表示部2822に、「操作メニュー1、再生2、プログラムアップロード3、読み上げ対象データダウンロード4、キャラクタ音声データダウンロード

ド・・」等のメッセージを表示する。ユーザが各操作を選択すると、記憶制御部 2821 は、通信処理部 2810 を通じて、記憶装置 2702 が公衆回線と接続されているか否かを判定し、接続されていない場合は、「回線と接続してください」等の接続を促すメッセージを記憶表示部 2822 に表示する。接続されている場合は、通信処理部 2810 を通じ、インターネット上のサーバ装置 2705 (図 27 参照) と接続する。

接続が完了すると、端末装置 2701 のシステム制御部 2801 は、操作の選択に従い、サーバ装置 2705 に読み上げ対象データ (該当する音声合成目的データ) やキャラクタ音声データ (該当する音素データ)、音声合成処理プログラムのダウンロードを要求し、サーバ装置 2705 から読み上げ対象データやキャラクタ音声データ、音声合成処理プログラムが送られてくると、各データを記憶装置 2702 に記憶する。そして、ダウンロード完了のメッセージを表示して、サーバ装置 2705 との呼を切断する。

なお、ユーザが記憶装置 2702 を端末装置 2701 に挿入して再生操作した場合の動作は、実施の形態 4 (図 12) で述べているので、その説明は省略する。

以上のように本実施の形態によれば、通信処理部 2810 を制御する記憶制御部 2821 を備え、記憶制御部 2821 は、通信処理部 2810 を介して、音声合成処理プログラムと音素データと音声合成目的データとをインターネット上のサーバ装置 2705 からダウンロードし、ダウンロードしたプログラムとデータとを端末装置インタフェース部 2820 と記憶装置インタフェース部 2803 とを介して記憶部 2805 に転送することにより、記憶装置単体でプログラムやデータをダウンロードすることができると共に、端末装置 2701 を簡単な構成にすることができ、端末装置 2701 のコストダウンを図ることができる。

(実施の形態 11)

図 29 は本実施の形態の携帯端末装置を有する通信システムを示す構成図である。図 29 において、要素 2901 ~ 2906 は図 201 ~ 206 と同一であり、説明は割愛する。肉声 2907 はユーザが本携帯端末装置に音声登録する。ユ

ユーザ独自で音声登録した音声で所望のテキストデータ等が朗読可能である。

図30は、図29の端末装置2901と記憶装置2902とを詳細に示すブロック図である。

要素3001～3010、3020と3022に関しては図1記載の要素101～110、120と122と同一であり、説明は割愛する。

マイク3011は音声登録する対象の肉声を入力する。音声入力処理部3012はアナログの肉声データをサンプリングしデジタルデータに変換する。音声登録処理部3013は音声入力処理部3012が変換したデジタル音声データを分析し、音素データベースを作成する。ユーザ登録の音素データベース3021は記憶装置2902内に記憶される。

このように構成された携帯端末装置について、その動作を説明する。
なお音声合成処理の動作については（実施の形態1）で述べているのでここでは割愛し、音声登録についての説明のみ行う。

図31は実施の形態1における携帯端末装置の音声登録処理部分のフローチャートである。ユーザが操作部3006から音声登録の操作を行うと（S3101）、システム制御部3001は音声入力処理部3012と音声登録処理部3013をアクティブにし、音声入力処理部3012はマイク3011から入力されるアナログの音声データをサンプリングしてデジタルデータに変換し、記憶部3005に記憶する（S3102）。

音声登録処理部3013は記憶部3005に記憶されている音声データを順次読み出し分析して、各音素毎に音の継続時間長や基本周波数、パワー等の音に関する情報やその音素が属するデータファイル名及びファイル中におけるその音素の開始位置や終了位置の情報等をラベル付けて、任意のフォーマットに沿った形でデータベース化して記憶装置2902に記録し、ユーザが端末装置を通じて入力した音声を登録する（S3103）。その後ユーザは音声キャラクタとして自分が登録した音声を選択し、所望の合成目的データを選択し再生操作することで合成音声出力される。この部分の動作は既に（実施の形態1）で述べているので割愛する。

このようにしてユーザは所望の音声キャラクタを自ら装置に登録でき、その音声キャラクタの音声にてテキストデータの朗読を聴くことができる。

(実施の形態 1 2)

- 5 図 3 2 は本実施の形態の携帯端末装置の詳細ブロック図である。3 2 3 0 は端末装置、3 2 4 0 は記憶装置である。要素 3 2 0 1 ~ 3 2 1 0、3 0 2 0 ~ 3 0 2 2 に関しては図 1 記載の要素 1 0 1 ~ 1 1 0、1 2 0 ~ 1 2 2 と同一であり、説明は割愛する。翻訳処理部 3 2 1 4 は元の合成目的データをユーザが所望する言語の合成目的データに変換する。

- 10 音声キャラクタの音素データベース及びテキストデータ等の合成目的データが記録された記憶装置 3 2 4 0 を端末装置 3 2 3 0 に挿入し、携帯端末装置上で翻訳の形態（例えば英語から日本語へ翻訳等）を選択し再生の操作を行うと携帯端末装置上からユーザが所望する翻訳形態及びキャラクタの音声でテキストデータの音声合成処理がなされ、音声合成音が出力される。

- 15 この携帯端末装置においてはユーザが再生操作を行うとまず合成目的データの翻訳処理が行われ、その翻訳データに関して音声合成処理が行われる。翻訳処理以外の動作は既に（実施の形態 1）で述べているので割愛する。

図 3 3 は実施の形態 1 2 における携帯端末装置の翻訳処理部分のフローチャートである。

- 20 ユーザが操作部 3 2 0 6 から再生の操作を行うと（S 3 3 0 1）、システム制御部 3 0 0 1 は表示部 3 2 0 7 に「朗読する言語を選択してください」等のようなメッセージと共に翻訳可能な言語を表示する。ユーザが操作部 3 2 0 6 を用いて朗読言語を選択する（S 3 3 0 2）と、記憶装置 I / F 部 3 2 0 3 に記憶装置 3 2 4 0 内にあると合成目的データ 3 2 2 2 を読み出す指令を出す。

- 25 記憶装置 I / F 部 3 2 0 3 は記憶装置 3 2 4 0 内にある携帯端末装置 I / F 部 3 2 2 0 とやり取りをしながら合成目的データ 3 2 2 2 を読み出し、携帯端末装置 3 2 3 0 内の記憶部 3 2 0 5 に保存する。

次にシステム制御部 3 2 0 1 は、翻訳処理部 3 2 1 4 に処理を開始させる。翻

訳処理部 3 2 1 4 はデータを解析し、ユーザが選択した言語の合成目的データに変換しながら、その変換データを記憶部 3 2 0 5 に記録する。その後音声合成処理部 3 2 0 2 が変換データを読み出し、(実施の形態 1) と同様な方法で合成処理を行い、合成音声出力される。

- 5 このようにしてユーザは所望する言語及びキャラクタ音声にてテキストデータ等の朗読を聴くことができる。

(実施の形態 1 3)

図 3 4 は本実施の形態の携帯端末装置の詳細ブロック図である。3 4 3 0 は端末装置、3 4 4 0 は記憶装置である。要素 3 4 0 1 ~ 3 4 1 0、3 4 2 0 と 3 4 2 2 に関しては(実施の形態 1) 図 1 記載の要素 1 0 1 ~ 1 1 0、1 2 0 と 1 2 2 と同一であり、説明は割愛する。要素 3 4 1 1 ~ 3 4 1 3 及び 3 4 2 1 は(実施の形態 1 1) の図 3 0 要素 3 0 1 1 ~ 3 0 1 3 及び 3 0 2 1 と、要素 3 4 1 4 は(実施の形態 1 2) 図 3 2 の要素 3 2 1 4 と同一であり、説明は割愛する。

- 15 この携帯端末装置においてはユーザが登録した音声キャラクタの音素データベース及びテキストデータ等の合成目的データが記録された記憶装置 3 2 4 0 を端末装置 3 2 3 0 に挿入し、携帯端末装置上で翻訳の形態(例えば英語から日本語へ翻訳等)を選択し再生の操作を行うと携帯端末装置上からユーザが所望する翻訳形態及びユーザが登録したキャラクタの音声での合成音声出力される。

- 20 音声登録処理の動作は(実施の形態 1 1) 図 3 1 にて説明しているので割愛する。

また翻訳処理の動作は(実施の形態 1 2) 図 3 3 にて説明しているので割愛する。またその他の動作は(実施の形態 1) 図 3 にて説明しているので割愛する。

- 25 ユーザは自身が登録した音声キャラクタ及び所望の言語にてテキストデータ等の朗読を聴くことができる。

(実施の形態 1 4)

図 3 5 は本実施の形態の携帯端末装置を有する通信システムを示す構成図及び

サーバ装置内のブロック図である。サーバ装置 3510 は合成目的データを供給する。なお説明の都合上、端末装置 3520 及び記憶装置 3530 の内部構成は実施の形態 1 図 1 と同様な構成とする。

ユーザは端末装置 3520 を通じ、自分が所望する小説等の合成目的データをダウンロードすることができる。その際合成目的データに関してはダウンロードするデータの範囲を選択可能である。

例えば小説等のデータであれば、その小説が書かれた日付や出版社名等が記述されたヘッダー情報や目次、及び本文等から構成され、ユーザは必ずしもヘッダー情報や目次まで朗読させようとは思わない可能性がある。このような場合に例えば「本文のみ」とかの選択が可能である。

サーバ装置 3510 ではユーザからのデータ範囲選択情報に基づき、合成目的データの構成分析を行い、指定範囲を抽出後その抽出した合成目的データをユーザに送る。送られた合成目的データは記憶装置 3530 内に記録され、ユーザが再生の操作を行うと携帯端末装置上からユーザが所望する合成目的データ範囲及びキャラクタの音声でテキストデータの音声合成処理がなされ、音声合成音が出力される。

次にサーバ装置内の詳細説明を行う。

サーバ装置のシステム制御部 3501 は装置内の各処理部とデータのやり取りを行い、装置全体の制御を行う。またネットワーク通信機能も含みインターネット上の通信も行う。サーバ装置の記憶部 3502 はサーバ装置の制御プログラムの記憶や様々な処理の作業領域等として用いられる。構文解析処理部 3503 はユーザから選択させたテキストデータの構成を分析し、ユーザが指定した部分のみ（例えば本文のみ等）のテキストデータを抽出する。合成目的データ保存部 3504 は複数の小説等テキストデータが保存されている。音素データベース保存部 3505 は複数のキャラクタの音素データベースが記憶されている。

次に動作の説明を行う。図 36 は実施の形態 1 4 におけるサーバ装置の動作フローチャートである。サーバ装置は先ず、ユーザからの接続要求を待つ（S3601）。ユーザがユーザ ID やパスワード等を用いて接続要求を行うと、サ

サーバー装置のシステム制御部は接続要求を行ったユーザが正規ユーザかどうか判断する（S3602）。正規ユーザでなかった場合は接続拒否をユーザに通知する（S3603）。正規ユーザであった場合は接続許可および合成目的データ保存部3504にある合成目的データのリスト情報を送信する。ユーザはこのリスト情報から合成目的データを選択、さらに選択した合成目的データのダウンロード選択部分（例えば本文のみ等）を選択する。サーバー装置はユーザからの合成目的データ及びダウンロード選択部分を示すデータの受信を待ち（S3605）、受信したら構文解析処理部3503が該当の合成目的データを合成目的データ保存部3504から読み出し、その構文を解析してユーザが選択したデータ部分を抽出する（S3606）。そして抽出したデータをユーザに配信する（S3607）。

ユーザは記憶装置に配信されたデータを記録し、端末装置に挿入、再生を行うことで合成音声出力される。このようにしてユーザは所望の音声キャラクタにて所望の範囲のテキスト等の朗読を聴くことができる。

（実施の形態15）

図37は本実施の形態の携帯端末装置を有する通信システムを示す構成図である。図37において、要素3701～3706は（実施の形態1）に記載した図2の要素201～206と同一である。楽譜データ3708はサーバー装置3705から携帯端末装置3701に提供される。

ユーザは端末装置3701を通じインターネットを経由してサーバー装置3705と接続し、聴きたい楽曲とその楽曲を歌わせたいキャラクタを選択する。するとサーバー装置3705からはユーザが選択した楽曲の歌詞に相当するテキスト等の合成目的データ3703と楽譜データ3708及び歌唱キャラクタの音楽データベース3704が送られてくる。送られてきたデータは端末装置装置経由で記憶装置に記憶され、ユーザが再生を行うとユーザが選択したキャラクタ音声で楽曲が再生される。

図38は携帯端末装置3701と記憶装置3702の詳細ブロック図である。

図38において、要素3801~3810及び3820~3822は（実施の形態1）の要素101~110及び120~122と同じであり説明は割愛する。

楽曲合成処理部3815は楽譜データを解析し、楽曲を構成する各音の高さや長さといった情報を解析する。楽譜データ3823は記憶装置内部に記憶される。

- 5 図39は実施の形態15における図38の携帯端末装置の動作フローチャートである。なお再生の動作以外は実施の形態1でのべた携帯端末装置と動作は同じであり、説明は割愛する。

- 操作部3806から再生が選択されると（S3901）、システム制御部3801は、表示部3807に楽曲データとキャラクタ音声のリストを表示し（S3902）、ユーザに楽曲データとキャラクタ音声データとを選択させる。ユーザが選択したら、記憶装置インタフェース部3803に、記憶装置3702内にある該当する楽曲データの楽譜データを読み出す指令を出し、記憶装置インタフェース部3803は記憶装置3702内にある端末装置インタフェース部3820とやり取りをしながら楽譜データを読み出し、端末装置内の記憶部3805に記録する。次に楽曲合成処理部3815は楽譜データを順次読み出し解析して楽曲を構成する各音の高さや長さといった音に関する情報を抽出する（S3803）。次にシステム制御部3801は記憶装置インタフェース部3803に、記憶装置3702内にある該当する楽曲データの歌詞データである合成目的データを読み出す指令を出し、データは記憶部3805に記録される。次に、システム制御部3801は、音声合成処理部3802に指令を出し、処理を開始させる。音声合成処理部3802は、記憶部3705から音声合成目的データを順次読み出しながら解析を行い、また楽曲合成処理部が抽出した音に関するデータを基に最も適する音素データを記憶装置3702から読み出して繋ぎ合わせ、楽曲データを作成し（S3904）、データを音声出力処理部3804が処理できるデータに変換して音声出力処理部3804に渡す。音声合成処理部3802は、ユーザが操作部3806で停止のボタンを押して、システム制御部3801から処理中止の指示があるまで、この処理を繰り返す。音声出力処理部3804は、音声合成処理部3802から受け取ったデータをフォーマット変換し、スピーカまたはヘッド

ドフォン3808に出力する(S3905)。このようにしてユーザは所望のキャラクター音声にて楽曲を聴くことができる。

(実施の形態16)

図40は本実施の形態の携帯端末装置のブロック図である。この装置は音声認証等で不正使用を防止するために合成目的データ中のある特定の文字列においてはユーザ所望のキャラクター音声ではなく代替の音声で朗読する携帯端末装置である

図40において、携帯端末装置は端末装置4030と記憶装置4040とを備える。要素4001~4010、4020~4022については(実施の形態1)の要素101~110及び120~122と同じであり説明は割愛する。テキスト解析処理部4016は合成目的データの中にお金の単位や数字等の特定の文字列が含まれていないか検索を行う。

図41は実施の形態16における図40の携帯端末装置の動作フローチャートである。なお再生の動作以外は実施の形態1でのべた携帯端末装置と動作は同じであり、説明は割愛する。

操作部4006から再生が選択されると(S4101)、システム制御部4001は、表示部4007に合成目的データとキャラクター音声のリストを表示し(S4102)、ユーザに合成対象データとキャラクター音声データとを選択させる。ユーザが選択したら、記憶装置インタフェース部4003に、記憶装置4040内にある該当する合成目的データを読み出す指令を出し、記憶装置インタフェース部4003は記憶装置4040内にある端末装置インタフェース部4020とやり取りをしながら合成目的データを読み出し、端末装置内の記憶部4005に記録する。次にテキスト解析処理部4016は合成目的データを順次読み出しながら文章の解析を行う。もしお金や数字等の特定文字列を発見した場合は、その文字列の先頭と最後に音声合成処理に影響の無い識別子を挿入しながら、記憶装置4040に記録する(S4103)。

次に、システム制御部4001は、音声合成処理部4002に指令を出し、処

音声合成処理部４００２はデータを解析しながらそのデータに最も適する音素データを記憶装置４０４０から読み出して繋ぎ合わせ、合成音声データを作成し（Ｓ４１０４）、データを音声出力処理部４００４が処理できるデータに変換して音声出力処理部４００４に渡す。音声合成処理部４００２は、ユーザが操作部４００６で停止のボタンを押して、システム制御部４００１から処理中止の指示があるまで、この処理を繰り返す。音声出力処理部４００４は、音声合成処理部４００２から受け取ったデータをフォーマット変換し、スピーカまたはヘッドフォン４００８に出力する（Ｓ３９０５）。

(実施の形態 17)

図42は実施の形態17における端末装置の動作フローチャートである。

なお再生の動作以外は実施の形態 1 でのべた携帯端末装置と動作は同じであり、説明は割愛する。

操作部から再生が選択されると（Ｓ４２０１）、システム制御部は、表示部に合成目的データとキャラクタ音声のリストを表示し（Ｓ４２０２）、ユーザに合成対象データとキャラクタ音声データとを選択させる。ユーザが選択したら、記憶装置インタフェース部に、記憶装置内にある該当する合成目的データを読み出

す指令を出し、記憶装置インタフェース部は記憶装置内にある端末装置インタフェース部とやり取りをしながら合成目的データを読み出し、端末装置内の記憶部に記録する。

次に、システム制御部は、音声合成処理部に指令を出し、処理を開始させる。

- 5 音声合成処理部は、記憶部から音声合成目的データを順次読み出しながら解析を行う。まず音声合成処理部は合成処理済の文字数を格納する変数 n を初期化し（S 4 2 0 4）、次に処理済みの文字数が識別音声挿入する文字数 T になったかを判断し（S 4 2 0 5）、そうである場合は n を 0 にして（S 4 2 0 6）、識別音データを合成音声データとして挿入する（S 4 2 0 7）。処理済みの文字数が
- 10 識別音声挿入する文字数 T に達していない場合は、処理を行おうとしている文字が句読点または句点を示すデータかを判断し（S 4 2 0 8）、そうである場合は識別音データを合成音声データとして挿入する（S 4 2 0 7）。そうでない場合はユーザが選択した音声キャラクタの音素データベースから最適な音素を抽出する（S 4 2 0 9）。

- 15 そして識別音データ及び音素データベースから抽出された音素データは順次接続され、合成音声データが作成される（S 4 2 1 0）。音声合成処理部は合成処理を行った文字数の変数をカウントアップし（S 4 2 1 1）、操作部から停止の指示があるまで（S 4 2 0 5）～（S 4 2 1 1）の処理が繰り返される。合成音声データはフォーマット変換され、スピーカやヘッドフォンに出力される（S 4
- 20 2 1 2）。

このようにこの装置においては音声認証等での不正使用を防止するために音声出力している文章の句点や句読点毎又は決められた文字数毎に音声合成の出力音であること示す音を強制的に挿入することができる。

25 (実施の形態 18)

図 4 3 は本実施の形態の携帯端末装置を有する通信システムを示す構成図である。携帯端末装置 4 3 0 1 は表示部や操作部、及びアンプやヘッドフォン、スピーカ等の音声出力手段を備える。メモリーカードや光ディスク及び磁気ディスク

等の記憶装置4302は音声合成目的のデータや音声キャラクタの音素データベース及び音声合成プログラムを格納し、携帯端末装置とは着脱可能である。

携帯端末装置における音声合成目的のデータ4303は、例えば小説等のテキストデータ等である。音素データベース4304は実在の人物の肉声をサンプリングし、そのサンプリングデータをデータベース化したものであり、この装置から出力される音声合成音の音色を決定する重要な要素となる。

インターネット上のサーバー装置4305は音声合成目的のデータ、例えば小説等のテキストデータや音素データベース及び音声合成プログラムを提供する。ここでは、一つのサーバー装置しか記載していないが、複数のサーバー装置で音声合成目的のデータと音素データベースを分けて提供する場合もある。4306はインターネット上のサーバー装置4305から提供され、携帯端末装置上で実行される音声合成処理プログラムである。

ユーザはまず音声キャラクタの音素データベースや音声合成目的のデータ及び音声合成処理プログラムを記憶している記憶装置4302を端末装置本体4301に挿入し、再生操作をすることにより、自分が所望するキャラクタの音素データベースを用いて音声合成処理を行い、あたかも当該キャラクタが発音したような音声にて音声合成目的のデータの朗読を聴くことができる。

または、ユーザはインターネット上のサーバー装置4305と接続し、自分が所望する小説等のテキストデータやキャラクタ音声を選択し、それらデータをサーバー装置4305からダウンロードして上で再生操作をすることにより、当該キャラクタが発音したような音声にて音声合成目的のデータの朗読を聴くことができる。この場合、サービス業者は音素提供者が発音した音声のサンプリングデータを分析して、音の強弱や音の高低等の音に関するデータの抽出を行い、データベース化してサーバー装置4305に保存している。

図44は本携帯端末装置のブロック図であり、図44において記憶装置4302は端末装置本体4301に接続可能な記憶装置である。端末装置本体4301において、システム制御部4401は端末装置本体4301に設けられ、装置内の各処理部とデータのやり取りを行い、装置全体の制御を行う。音声合成処理部

4302は音声合成目的のデータの解析を行って、合成目的のデータ（各文字データまたは単語のデータ）に最適な音素データを抽出及び連結する。記憶装置I/F部4403は記憶装置4302へのデータの書き込み、データの読み出しをする。

- 5 音声出力手段4404は音声合成処理部4402からのデータを受け、データのフォーマット変換等の処理を行い、デジタルデータをアナログデータに変換し、不要なノイズをカットしてスピーカまたはヘッドフォン等よりに出力する。携帯端末装置の記憶部4405は携帯端末装置を制御するプログラムや表示部に表示するフォントデータの記憶、音声合成音データの保存及び様々なデータ処理の際の作業領域等として用いられる。

操作部4406を通じユーザは装置に自分の指示を与える。表示部4407は装置の動作状態等をユーザに表示する。電源部4408は装置に電源を供給する。通信処理部4409は公衆回線と接続して、インターネット上でのデータのやり取りを実現する。

- 15 音素データベース選択手段4410は音声合成目的のデータを解析し、音声合成処理に用いる音素データベースを選択する。端末装置I/F部4420は記憶装置I/F部4403と共に端末装置4301とのデータのやり取りを行う。

次に記憶装置4302は音素データベース4421と音声合成目的のデータ4422を保存する。なお、記憶装置4302内及び端末装置の記憶部に保存されるのは、音声合成目的のデータや音素データベース及び音声合成処理プログラムのどのような組み合わせであっても良い。

次に、本発明の携帯端末装置の動作について図45の動作フローチャートを用いて説明する。ユーザが操作部にて装置の電源をオンすると、システム制御部4401は記憶装置I/F部4403に指令を出し、記憶装置4302が端末装置本体4401に接続しているか調べる（S4501）。接続していない場合は表示部4407に「メモリーカードを挿入してください。」等のユーザに対して端末装置本体4301に記憶装置4302の接続を行うように促すメッセージを表示する（S4502）。

記憶装置4302が端末装置本体4301に接続している場合は「操作メニュー 1. 朗読 2. プログラムアップデート 3. 音声合成目的データのダウンロード 4. キャラクタ音声データダウンロード・・・」等のユーザに端末の操作を促す操作メニューを表示部4407に表示する(S4503)。

5 ユーザーが操作部4406を操作し、朗読以外の操作が選ばれたら(S4504)、システム制御部4401は通信処理部を通じて端末装置本体4301が公衆回線と接続されているかを調べ(S4505)、接続していない場合は「回線と接続してください」等の回線の接続を促すメッセージを表示部4307に表示する(S4506)。

10 端末装置本体4301が公衆回線に接続している場合は、通信処理部4409を介し、インターネット上のサーバー装置4305に接続する(S4507)。端末装置本体4301のシステム制御部4401はサーバー装置4305に音声合成プログラムや音声合成目的のデータや音素データデータベースのダウンロードを要求し(S4508)、サーバー装置4305から前記データが送られてく
15 ると、データを記憶部4405や記憶装置4302に記憶する(S4509)。それが完了すると、サーバー装置4305との呼を切断する。

前述の操作メニューが表示されている状態でユーザーが操作部4406を操作して朗読が選択されると(S4504)、システム制御部4401は表示部4407に装置内に存在する音声合成目的のデータのリスト、および音声キャラクタ
20 のリストを表示する。ユーザは操作部4406を操作して音声合成目的のデータを決定する(S4510)と、音素データベース選択手段4410は音声合成目的のデータを解析して、各々の音素データベースを適用するデータの範囲を抽出する(S4511)。例えば音声合成目的のデータが小説のテキストデータの場合は、登場人物のセリフ部分やナレーション部分等にデータ範囲を分け、その結
25 果をシステム制御部に伝える。

システム制御部4401は音素データベース選択手段4410からの結果を基に表示部4407に「次の部分に適用する音声キャラクタを選択してください。

1. 登場人物Aセリフ 2. 登場人物Cセリフ 4. ナレーション」等のような

音声合成目的データの前記各部分の読み上げに割り当てる音声キャラクタを選択させるメッセージを表示させる。

ユーザは操作部4406を操作して音声キャラクタを決定する(S4512)。ユーザは場合によっては複数の人物を指定することが可能であり、例えば小説の

中の複数の登場人物毎に音声キャラクタを変えて指定することもある。

システム制御部4401は選択された結果を音素データベース選択手段4410に伝える。音素データベース選択手段4410はこの結果を基に音声合成目的データの各音素データベースを適用する部分に対して識別記号を混在させ(S4513)、音声合成処理部4402が音声合成目的データ中のどの部分でどの音声キャラクタの音素データベースを使用すればよいのかを判別できるようにして結果を記憶部4405に記憶する。すなわち、音声合成目的データの中で部分毎に適切な音声キャラクタを示す識別記号が加えられる。

これにより、音声合成処理時には、音声合成処理部4402は音声合成目的データの中の部分ごとに適切な音声キャラクタの音素データベースを使用して音声合成を行い、例えば小説であれば登場人物のセリフ毎にキャラクタを変えて音声合成することができ、よりリアルな読み上げを実現することができる。なおこの音素データベース選択手段において、各音素データベースを適用させるデータ範囲の分け方は前記のような登場人物のセリフ毎であったり、章毎であったり、行毎であったりしても良く、その分け方は音声合成目的のデータ内容にも依存するので限定はしない。

次にシステム制御部4401は音声合成処理部4402に処理を開始させる。音声合成処理部4402は記憶部4405から音素データベース選択手段が処理したデータを順次読み出し、識別記号に基づき使用する音声キャラクタの音素データベースを選択する。そして音声合成目的のデータを分析し、各データに最も適する音素データを記憶部4405または記憶装置4302から読み出して、繋ぎ合わせ合成音声データを作成する(S4514)。

音声出力手段4404は音声合成処理部4402から合成音声データを受け取り、データをフォーマット変換して、スピーカまたはヘッドフォンに出力する

(S 4 5 1 5)。

なお、データ入力手段としてメモリーカードや光ディスク等の記憶装置 4 3 0 2 を用いたものであるが、データ入力手段がモデム等のネットワークインターフェース及びキーボードであっても良い。また通信処理部 4 4 0 9 が端末装置本体 5 (2 0 1) に搭載されたものであるが、通信処理部を記憶装置 4 3 0 2 に搭載し、ネットワーク上にあるサーバー装置から音素データベースや音声合成目的のデータ及び音声合成処理プログラムをダウンロードして記憶装置に記憶するようにしても良い。

また、音声合成目的のデータとして、小説等のテキストデータのみでなく、スコアデータ (楽譜) + テキストデータ (歌詞) の楽曲データとし、ユーザが端末装置本体 4 3 0 1 を操作して所望のキャラクタと楽曲データを指定することで、当該キャラクタの音素データベースと前述の楽曲データを基に音声合成処理を行い、ユーザが所望する特定キャラクタの音声で楽曲を聴くこともできる。

15 (実施の形態 1 9)

図 4 6 は本実施の形態の携帯端末装置を有する通信システムを示す構成図である。携帯端末装置 4 6 0 1 は表示部や操作部、及びアンプやヘッドフォン、スピーカ等の音声出力手段を備える。メモリーカードや光ディスク及び磁気ディスク等の記憶装置 4 6 0 2 は音声合成目的のデータや音声キャラクタの音素データベース及び音声合成プログラムや挿絵等の画像データを格納し、携帯端末装置とは着脱可能である。

携帯端末装置における音声合成目的のデータ 4 6 0 3 は、例えば小説等のテキストデータ等である。音素データベース 4 6 0 4 は実在の人物の肉声をサンプリングし、そのサンプリングデータをデータベース化したものであり、この装置から出力される音声合成音の音色を決定する重要な要素となる。

インターネット上のサーバー装置 4 6 0 5 (2 0 5) は音声合成目的のデータ、例えば小説等のテキストデータや音素データベース及び音声合成プログラムや画像データを提供する。ここでは、一つのサーバー装置しか記載していないが、複

数のサーバー装置で各データを分けて提供する場合もある。

音声合成処理プログラム4606はインターネット上のサーバー装置4605から提供され、携帯端末装置上で実行される。画像データ4607は音声合成目的のデータに関連する挿絵等のデータである。

5 ユーザはまず音声キャラクタの音素データベースや音声合成目的のデータ及び音声合成処理プログラム及び画像データを記憶している記憶装置4602を端末装置本体4601挿入し、再生操作をすることにより、自分が所望するキャラクタの音素データベースを用いて音声合成処理を行い、あたかも当該キャラクタが発音したような音声にて音声合成目的のデータの朗読を聴くことができる。

10 または、ユーザはインターネット上のサーバー装置4605と接続し、自分が所望する小説等のテキストデータやキャラクタ音声を選択し、それらデータをサーバー装置4605からダウンロードして上で再生操作をすることにより、当該キャラクタが発音したような音声にて音声合成目的のデータの朗読を聴くことができる。この場合、サービス業者は音素提供者が発音した音声のサンプリングデータ

15 データを分析して、音の強弱や音の高低等の音に関するデータの抽出を行い、データベース化してサーバー装置4605に保存している。

図47は本携帯端末装置のブロック図である。図47において記憶装置4602は端末装置本体4601に接続可能な記憶装置である。

端末装置本体4601において、端末装置本体4601に設けられたシステム

20 制御部4701は装置内の各処理部とデータのやり取りを行い、装置全体の制御を行う。音声合成処理部4702は音声合成目的のデータの解析を行って、合成目的のデータ（各文字データまたは単語のデータ）に最適な音素データを抽出及び連結する。記憶装置I/F部4703は記憶装置4602へのデータの書き込み、データの読み出しをする。

25 音声出力手段4704は音声合成処理部4702からのデータを受け、データのフォーマット変換等の処理を行い、デジタルデータをアナログデータに変換し、不要なノイズをカットしてスピーカまたはヘッドフォン等より出力する。携帯端末装置の記憶部4705は携帯端末装置を制御するプログラムや表示部470

7に表示するフォントデータの記憶、音声合成音データの保存及び様々なデータ処理の際の作業領域等として用いられる。

操作部4706を通じユーザは装置に自分の指示を与える。表示部4707は装置の動作状態や読み上げ中の音声合成目的データ及び挿絵等の画像データ等をユーザに表示する。電源部4708は装置に電源を供給する。通信処理部4709は公衆回線と接続して、インターネット上でのデータのやり取りを実現する。

文字データ表示処理部4710は、合成音声出力中の音声合成目的データの表示を行う。画像データ表示処理部4711は、合成音声出力中の箇所の内容に合った挿絵や画像の表示を行う。端末装置I/F部4720は記憶装置I/F部4703と共に端末装置4602とのデータのやり取りを行う。

次に記憶装置4602は音素データベース4721と音声合成目的のデータ4722を保存する。画像データ4723は、音声合成目的データに関連する挿絵等のデータである。なお、記憶装置4602内及び端末装置の記憶部に保存されるのは、音声合成目的のデータや音素データベース及び音声合成処理プログラム及び画像データのどのような組み合わせであっても良い。

次に、本発明の携帯端末装置の動作について図48の動作フローチャートを用いて説明する。ユーザが操作部にて装置の電源をオンすると、システム制御部4701は記憶装置I/F部4703に指令を出し、記憶装置4602が端末装置本体4601に接続しているか調べる(S4801)。接続していない場合は表示部4707に「メモリーカードを挿入してください。」等のユーザに対して端末装置本体4601に記憶装置4602の接続を行うように促すメッセージを表示する(S4802)。

記憶装置4602が端末装置本体4601に接続している場合は「操作メニュー 1. 朗読 2. プログラムアップデート 3. 音声合成目的データのダウンロード 4. キャラクタ音声データダウンロード 5. 画像データダウンロード・・・」等のユーザに端末の操作を促す操作メニューを表示部4707に表示する(S4803)。システム制御部4701は操作部4706からの操作状況を監視している。

ユーザーが操作部4706を操作し、朗読以外の操作が選ばれたら（S4804）、システム制御部4701は通信処理部を通じて端末装置本体4601が公衆回線と接続されているかを調べ（S4805）、接続していない場合は「回線と接続してください」等の回線の接続を促すメッセージを表示部4707に表示する（S4806）。

端末装置本体4601が公衆回線に接続している場合は、通信処理部4709を介し、インターネット上のサーバー装置4605に接続する（S4807）。端末装置本体4601のシステム制御部4701はサーバー装置4605に音声合成プログラムや音声合成目的のデータや音素データデータベースや画像データのダウンロードを要求し（S4508）、サーバー装置4605から前記データが送られてくると、データを記憶部4705や記憶装置4602に記憶する（S4809）。それが完了すると、サーバー装置4605との呼を切断する。

操作部4706から朗読が選択されると（S4804）、システム制御部4701は表示部4707に装置内に存在する音声合成目的データのリスト、および音声キャラクタのリストを表示する。ユーザは操作部4706を操作して音声合成目的のデータを決定する（S4810）と、文字データ表示処理部4710は音声合成目的データ4722を記憶部4705または記憶装置4602から読み出し、表示部4707の表示領域内に表示可能な文字数の範囲内で文字表示識別符号を合成目的のデータに混在して記憶部4705または記憶装置4602に記録する（S4810）。この文字表示識別符号は合成音声出力中の合成目的データ、例えば文字を表示部4707に表示させる為のものである。

文字表示識別符号の合成目的データへの混在の方法は複数あり、表示する文字列の最初の文字データから最後の文字データを囲んで混在させたり、表示する文字列の先頭の文字のみに混在させる方法等があり、限定はしない。

次に音声合成処理部4702は文字データ表示処理部が処理したデータを読み出し、場合によっては音声合成対象のデータを音声合成が可能なデータに変換する。そして合成解析対象のデータが画像表示識別符号かをチェックする（S4811）。この画像表示識別符号は合成音声出力中の合成目的データの内容に合致

した、或いはその内容理解を助ける為の画像データを表示する為のものである。
画像表示識別符号は合成目的データに予め混在されても良いし、画像データ表示
処理部4711が合成目的データを解析して、合成目的データの内容に合う画像
データの識別符号を混在させても良い。

- 5 もし画像表示識別符号であった場合は、その符号内容を画像データ表示処理部
4711に伝え、画像データ表示処理部4711は記憶部4705または記憶装
置4602から符号内容に対応する画像データ読み出し、表示部4707に表示
する(S4812)。

- 10 画像表示識別符号でなかった場合、音声合成処理部4702は解析対象データ
が文字表示識別符号かをチェックする(S4813)。文字表示識別符号であっ
た場合は、その符号を文字データ表示処理部4710に伝え、文字データ表示処
理部4710は伝えられた符号で囲まれた部分の文字列データを表示部4707
に表示したり、あるいは同じ符号がついた先頭の文字から表示領域に表示可能な
範囲での決められた文字数分の文字を表示部4707に表示する(S4814)。

- 15 そして解析対象データが文字表示識別符号でなかった場合、音声合成処理部4
702の解析対象は実際の音声合成対象データとなり、そのデータに最も適する
音素データを記憶部4705または記憶装置4702から読み出して、繋ぎ合わ
せ合成音声データを作成する(s315)。

- 20 音声出力手段は音声合成処理部4702が作成した合成音声データを受け取り、
データをフォーマット変換して、スピーカ又はヘッドフォンから合成音声を出力
する(S4816)。

- 25 なお、データ入力手段としてメモリーカードや光ディスク等の記憶装置460
2を用いたものであるが、データ入力手段がモデム等のネットワークインターフ
ェース及びキーボードであっても良い。また通信処理部4709が端末装置本体
4601に搭載されたものであるが、通信処理部を記憶装置4702に搭載し、
ネットワーク上にあるサーバー装置から音素データベースや音声合成目的のデー
タ及び音声合成処理プログラムや画像データをダウンロードして記憶装置460
2に記憶するようにしても良い。

(実施の形態 20)

図 49 は本実施の形態の読上げシステムを示す構成図である。携帯端末装置 4901 は表示部や操作部及びヘッドフォン、スピーカ等の音声出力手段を備える。
5 ユーザが選択したキャラクタ音声、及び合成目的データでの合成音声データ 4902 はサーバー装置 4903 からユーザに提供される。

ユーザはまず携帯端末装置 4901 を通じてインターネット上のサーバー装置 4903 と接続し、自分が所望する小説のテキストデータ等である合成目的データやキャラクタ音声を選択する。サーバー装置 4903 は選択された合成目的データ及びキャラクタの音素データベースを用いて音声合成処理を行い、合成音声データをインターネット経由でユーザに送信し、ユーザは携帯端末装置内にそのデータを記憶する。そして再生の処理を行うと携帯端末装置からユーザが所望するキャラクタの音声でテキストデータの音声合成音出力される。

図 50 は携帯端末装置 4901 とサーバー装置 4903 の詳細説明図である。
15 図 50 において、携帯端末装置のシステム制御部 5001 は装置内の各処理部とデータのやり取りを行い、装置全体の制御を行う。またネットワーク通信機能も含みインターネット上の通信も行う。携帯端末装置の記憶部 5002 は携帯端末装置の制御プログラムや表示部に表示するフォントデータの記憶、サーバー装置 4903 から送られる音声合成音データの保存及び様々な処理の作業領域等として用いられる。音声出力処理部 5003 はデジタル音声データをアナログデータに変換し、不要なノイズをカットしてまたスピーカ 5004 に出力する。操作部 5005 を通じユーザは装置に自分の指示を与える。表示部 5006 は装置の動作状態等をユーザに表示する。電源部 5007 は装置に電源を供給する。

サーバー装置のシステム制御部 5021 は装置内の各処理部とデータのやり取りを行い、装置全体の制御を行う。またネットワーク通信機能も含みインターネット上の通信も行う。サーバー装置の記憶部 5022 はサーバー装置の制御プログラムの記憶や様々な処理の作業領域等として用いられる。音声合成処理部 5023 は合成目的データの解析を行って、各文字データに最適な音素データを抽出

及び連結してそのデータを携帯端末装置内にある音声出力処理部5003に渡せるようにデータ変換する。合成目的データ保存部5024は複数の合成目的データ、例えば小説等の文章データが保存されている。音素データベース保存部5025は複数のキャラクタの音素データベースが記憶されている。

5 以下に本読み上げシステムにおける詳細な動作説明を行う。まずユーザがサーバー装置から合成音声データをダウンロードする場合の内部動作の説明を行う。

図51は本読み上げシステムの動作フローチャートである。サーバー装置は先ず、ユーザからの接続要求を待つ(S5101)。ユーザがユーザIDやパスワード等を用いて接続要求を行うと、サーバー装置のシステム制御部5021は接続要求を行ったユーザが正規ユーザかどうか判断する(S5102)。

10 正規ユーザでなかった場合は接続拒否をユーザに通知する(S5103)。正規ユーザであった場合は接続許可および合成目的データ保存部5024にある合成目的データと音素データベース保存部5025にある音声キャラクタのリスト情報を送信する(S5104)。次にユーザが選択した所望の合成目的データと音声キャラクタ

15 データのリクエストデータを受信すると(S5105)、サーバー装置のシステム制御部5021は音声合成処理部5023に処理を開始させる。音声合成処理部5023はサーバー装置の記憶部5022から合成目的データを順次読み出しながら解析を行い、各文字データに最も適する音素データを記憶部5022から読み出して、繋ぎ合わせ、合成音声データを作成する(S5106)。そして合

20 成音声データをユーザに送信する(S5107)。このようにしてユーザは所望する合成目的及びキャラクタ音声の合成音データを得ることができる。

次に携帯端末装置側での合成音の再生について述べる。合成音声データがすべて携帯端末装置の記憶部5002に記憶されると、システム制御部5001は操作部5005から再生ボタンが押されるのを待ち、押されたら、音声出力処理部

25 5003をアクティブにする。音声出力処理部5003は記憶部5002から次々と合成音声データを読み出し、アナログ音声に変換してスピーカ5004に音声出力する。

このようにしてユーザは所望するキャラクタ音声にてテキストデータ等の朗読

を聴くことができる。

(実施の形態 21)

図 52 は本実施の形態の読上げシステムを示す構成図である。要素 5201 ~ 5203 は (実施の形態 20) に記載した図 49 の要素 4901 ~ 4903 と同一である。肉声 5204 はユーザが音声登録する。肉声サンプリングデータ 5205 は携帯端末装置内に取り込まれる。

ユーザはまず携帯端末装置 5201 を用いて、サーバー装置 5203 に登録したい肉声 5204 を携帯端末装置 5201 に取り込む。次にサーバー装置 5203 とインターネット経由で接続し、肉声サンプリングデータ 5205 をサーバー装置 5203 に送る。サーバー装置内部では音声登録処理が行われ、肉声サンプリングデータ 5205 が音素データベースとして登録される。そして、携帯端末装置 5201 を通じて所望の合成目的データとキャラクタ音声として先に登録したものを選択する。サーバー装置 5203 からはユーザの選択に基づき、処理を行い合成音声データ 5202 をユーザの携帯端末装置 5201 に送信する。ユーザは携帯端末装置 5201 に合成音声データ 5202 を取り込み、再生を行うことにより、携帯端末装置からはユーザが選択した合成目的データ及び、自ら登録したキャラクタ音声でテキストデータ等の朗読を聴くことができる。

図 53 は携帯端末装置 5201 とサーバー装置 5203 の詳細説明図である。

図 53 において、5301 ~ 5307 は (実施の形態 20) に記載した図 50 の 5301 ~ 5307 と同一である。5309 はユーザが音声登録するための肉声を入力するためのマイク、5308 はマイクから取り込まれた肉声のアナログデータをサンプリングし、デジタルデータに変換する音声入力処理部である。5321 ~ 5325 は (実施の形態 20) に記載した図 50 の 5021 ~ 5025 と同一である。5326 は携帯端末装置 5201 の音声入力処理部 5308 が変換したデジタル音声データを分析し、音素データベースを作成する音声登録処理部である。

以下に本実施の形態の読上げシステムにおける詳細な動作説明を行う。図 5

4は同読み上げシステムの動作フローチャートである。ここでは音声の登録処理までの動作について述べる。ユーザが合成音声データをサーバー装置から得るまでの動作及び端末装置における音声データの再生は実施の形態20の図51で述べたとおりであり、割愛する。

5 ユーザが携帯端末装置5201の操作部5305から音声登録の操作を行うと（S5401）、携帯端末装置内のシステム制御部5301は音声入力処理部5308をアクティブにし、マイク5309から入力されるアナログの音声データをサンプリングしてデジタルデータに変換し（S5402）、携帯端末装置内の記憶部5302に記憶する。サーバー装置5203はユーザからの接続要求を待ち（S5121）、ユーザがユーザIDやパスワード等を用いて接続要求を行うと（S5403）、サーバー装置5203のシステム制御部5301は接続要求を行ったユーザが正規ユーザかどうか判断する（S5422）正規ユーザでなかった場合は接続拒否をユーザに通知する（S5423）。正規ユーザであった場合は接続許可を通知する（S5424）。端末装置5201はサーバー装置5203から接続許可のメッセージを受けると（S5404）、音声登録の依頼をする為に肉声のサンプリングデータを記憶部5302から読み出し、サーバー装置5203に送信する（S5405）。サーバー装置5203はユーザからのデータを受信すると（S5425）、受信データを記憶部5322に記憶する。次に音声登録処理部5326は記憶部5322に記憶されている音声データを順次読み出し分析して、各音素毎に音の継続時間長や基本周波数、パワー等の音に関する情報やその音素が属するデータファイル名及びファイル中におけるその音素の開始位置や終了位置の情報等をラベル付けして、任意のフォーマットに沿った形でデータベース化して記憶装置2902に記録し、ユーザが端末装置を通じて入力した音声に登録する（S5426）。そしてユーザに登録完了通知を送信する（S5427）。端末装置ではサーバー装置からの音声登録の完了通知を待ち、待っている間、表示部5306に「音声登録処理中です。」等のメッセージを表示する（S5406、S5407）。サーバー装置から登録の完了通知がくると端末装置のシステム制御部（S5301）は登録の完了メッセージを表示部53

06に表示する(S5408)。

ユーザは登録した音声を読ませるキャラクタ音声として選択すれば、前記実施の形態20で述べたような同様の動作にて、携帯端末装置から音声合成音出力される。このようにしてユーザは所望するテキストデータを自ら登録したキャラクタ音声中で聴くことができる。

(実施の形態22)

図55は本実施の形態の読上げシステムを示す構成図である。要素5501～5503は(実施の形態20)に記載した図49の要素4901～4903と同一である。

ユーザは携帯端末装置5501を通じてインターネット上のサーバー装置5503と接続し、自分が所望する小説等の合成目的データや朗読の言語及びキャラクタ音声を選択する。サーバー装置5503は選択された言語と合成目的データ及びキャラクタの音素データベースを用いて音声合成処理を行い、合成音声をインターネット経由でユーザに送信し、ユーザは携帯端末装置内にそのデータを記憶する。そして再生の処理を行うと携帯端末装置からユーザが指定した言語及びキャラクタの音声中で合成目的データの合成音声が出力される。このようにしてユーザは所望する言語及びキャラクタ音声中でテキストデータ等の朗読を聴くことができる。

図56は携帯端末装置5501とサーバー装置5503の詳細説明図である。図56において、要素5601～5607は(実施の形態20)の図50の要素5001～5007と同一である。また要素5621～5625は(実施の形態20)の図50の要素5021～5025と同一である。また5627は元の合成目的データをユーザが所望する言語のデータに変換する翻訳処理部である。

以下に本実施の形態の読上げシステムにおける詳細な動作説明を行う。図57は同読上げシステムの動作フローチャートである。サーバー装置は先ず、ユーザからの接続要求を待つ(S5701)。ユーザがユーザIDやパスワード等を用いて接続要求を行うと、サーバー装置のシステム制御部5621は接続要求

を行ったユーザが正規ユーザかどうか判断する（S5702）。正規ユーザでなかった場合は接続拒否をユーザに通知する（S5703）。正規ユーザであった場合は接続許可および合成目的データ保存部5624にある合成目的データと音素データベース保存部5625にある音声キャラクタ及び翻訳可能な言語のリスト情報を送信する（S5704）。次にユーザが選択した所望の合成目的データと音声キャラクタデータ及び翻訳言語のリクエストデータを受信すると（S5705）、サーバー装置のシステム制御部5621は、翻訳処理部5627に処理を開始させる。翻訳処理部5627はデータを解析し、ユーザが選択した言語の合成目的データに変換しながら、その変換データを記憶部5622に記録する（S5606）。その後サーバー装置のシステム制御部5621は、音声合成処理部5623に処理を開始させる。音声合成処理部5623はサーバー装置の記憶部5622から合成目的データを順次読み出しながら解析を行い、各文字データに最も適する音素データを記憶部5622から読み出して、繋ぎ合わせ、合成音声データを作成する（S5707）。そして合成音声データをユーザに送信する（S5107）。このようにしてユーザは所望する合成目的及びキャラクタ音声の合成音データを得ることができる。

ユーザは送られた合成音声データを端末装置に取りこみ再生することで端末装置からは合成音声が出力される。

このようにしてユーザは所望の音声キャラクタ及び言語にてテキストデータ等の朗読を聴く事ができる。

（実施の形態23）

図58は本実施の形態の読上げシステムを示す構成図である。要素5801～（5804）は（実施の形態21）に記載した図52の要素5201～5204と同一である。

ユーザは（実施の形態21）に述べたようにして登録したいキャラクタの音声に登録する。次に（実施の形態22）に述べたような方法で自分が所望する小説等の合成目的データや朗読の言語及びユーザが登録したキャラクタ音声を選択す

る。サーバー装置5203は選択された言語と合成目的データ及びユーザ登録のキャラクタの音素データベースを用いて音声合成処理を行い、その出力結果をインターネット経由でユーザに送信し、ユーザは携帯端末装置内にそのデータを記憶する。そして再生の処理を行うと携帯端末装置からユーザが指定した言語及びユーザ登録のキャラクタの音声で合成目的データの合成音声出力される。このようにしてユーザは所望する言語及び登録したキャラクタ音声にてテキストデータ等の朗読を聴くことができる。

図59は携帯端末装置5801とサーバー装置5803の詳細説明図である。

図59において、要素5901～5909は（実施の形態21）の図53の要素5301～5309と同一である。また要素5921～5926は（実施の形態21）の図53の要素5321～5326と同一である。また要素5927は（実施の形態22）の図56の要素5627と同一である。

携帯端末装置5801とサーバー装置5803における音声登録の内部動作は（実施の形態22）に記載したものと同一であり、また翻訳処理の内部動作は（実施の形態23）に記載したものと同一であるので説明は割愛する。ユーザは所望する言語及び登録したキャラクタ音声にてテキストデータ等の朗読を聴くことができる。

（実施の形態24）

図60は本実施の形態の読上げシステムを示す構成図である。図60において、要素6001、6003は（実施の形態20）に記載した図49の要素4901、4903と同一である。6002はサーバー装置6003から携帯端末装置6001に提供される楽曲データである。

ユーザはまずインターネットを経由してサーバー装置6003と接続し、聴きたい楽曲とその楽曲を歌わせたい音声キャラクタを選択する。サーバー装置6003ではユーザの指示に従い、楽譜データを解析し、解析された音に関するデータを基に選択された楽曲の歌詞データと選択されたキャラクタの音素データベースを用いて音声合成処理し、そのデータをインターネットを経由してユーザに提

供する。ユーザは携帯端末装置 6 0 0 1 にその楽曲データを取り込み再生することで、ユーザが所望するキャラクタ音声にて楽曲を聴く事ができる。

図 6 1 は携帯端末装置 6 0 0 1 とサーバー装置 6 0 0 3 の詳細説明図である。図 6 0 において、携帯端末装置側の 6 1 0 1 ~ 6 1 0 7 は（実施の形態 2 1）の図 5 0 の 5 0 0 1 ~ 5 0 0 7 と同一である。またサーバー装置側の 6 1 2 1 ~ 6 1 2 5 は図 5 0 の 5 0 2 1 ~ 5 0 2 5 と同一である。6 1 2 6 は楽曲の楽譜データを保存する楽譜データ保存部である。6 1 2 7 は楽譜データを解析し、楽曲を構成する各音の高さや長さといった情報を解析する楽曲合成処理部である。

図 6 2 は本実施の形態の読み上げシステムの動作フローチャートである。サーバー装置は先ず、ユーザからの接続要求を待つ（S 6 2 0 1）。ユーザがユーザ ID やパスワード等を用いて接続要求を行うと、サーバー装置のシステム制御部 6 1 2 1 は接続要求を行ったユーザが正規ユーザかどうか判断する（S 6 2 0 2）。正規ユーザでなかった場合は接続拒否をユーザに通知する（S 6 2 0 3）。正規ユーザであった場合は接続許可および合成目的データ保存部 6 1 2 4 にある合成目的データ（歌詞データ）と音素データベース保存部 6 1 2 5 にある音声キャラクタのリスト情報を送信する（S 6 2 0 4）。次にユーザが選択した所望の楽曲データと音声キャラクタデータのリクエストデータを受信すると（S 6 2 0 5）、サーバー装置のシステム制御部 6 1 2 1 は該当の楽譜データを楽譜データ保存部 6 1 2 6 から読み出し、サーバー装置内の記憶部 6 1 2 2 に記録する。次に楽曲合成処理部 6 1 2 7 は楽譜データを順次読み出し解析して楽曲を構成する各音の高さや長さといった音に関する情報を抽出する（S 6 2 0 6）。次にシステム制御部 6 1 2 1 は該当する楽曲データの歌詞データである合成目的データを読み出す指令を出し、データは記憶部 6 1 2 2 に記録される。次に、システム制御部 6 1 2 1 は、音声合成処理部 6 1 2 3 に指令を出し、処理を開始させる。音声合成処理部 6 1 2 3 は、記憶部 6 1 2 2 から音声合成目的データを順次読み出しながら解析を行い、また楽曲合成処理部が抽出した音に関するデータを基に最も適する音素データを音素データベース保存部 6 1 2 5 から読み出して繋ぎ合わせ、合成音声データ（楽曲データ）を作成する（S 6 2 0 7）。そしてこの合成

音声データをユーザに送信する（S6208）。

このようにしてユーザは所望する合成目的及びキャラクタ音声の合成音データを得ることができる。そしてこの合成音声データを端末装置に取りこみ再生することで、合成音声（楽曲データ）が再生される。このようにしてユーザは所望の

5 キャラクタ音声で楽曲を聴く事ができる。

（実施の形態25）

図63は本実施の形態の読上げシステムを示す構成図である。図63において要素6301～6303は（実施の形態20）における図49の要素4901、

10 4903と同一である。この読上げシステムにおける携帯端末装置は出力される音声は音声認証等での不正使用を防止するために合成目的データ中のある特定の文字列においてはユーザ所望のキャラクタ音声ではなく代替の音声で出力される。

図64は携帯端末装置6301とサーバー装置6303の詳細説明図である。要素6401～6407は（実施の形態20）の図49の要素4901～490

15 7と同一である。またサーバー装置側の要素6421～6425は図49の要素4921～4925と同一である。テキスト解析処理部6426は合成目的データの中にお金の単位や数字が含まれていないか検索を行う。

次に本実施の形態の読上げシステムの動作説明を行う。図65は同読上げシステムの動作フローチャートである。サーバー装置はまず、ユーザからの接続

20 要求を待つ（S6501）。ユーザがユーザIDやパスワード等を用いて接続要求を行うと、サーバー装置のシステム制御部6421は接続要求を行ったユーザが正規ユーザかどうか判断する（S6502）。正規ユーザでなかった場合は接続拒否をユーザに通知する（S6503）。正規ユーザであった場合は接続許可および合成目的データ保存部6424にある合成目的データと音素データベース

25 保存部6425にある音声キャラクタのリスト情報を送信する（S6504）。次にユーザが選択した所望の合成目的データと音声キャラクタデータのリクエストデータを受信すると（S6505）、サーバー装置のシステム制御部6421は該当の合成目的データを合成目的データベース6424から読み出し、サーバ

一装置の記憶部 6 4 2 2 に記憶する。

次にテキスト解析処理部 6 4 2 6 は合成目的データを記憶部 6 4 2 2 から順次読み出して文章の解析を行う。もしお金や数字等の特定文字列を発見した場合は、その文字列の先頭と最後に音声合成処理に影響の無い識別子を挿入しながら、記憶部 6 4 2 2 に記録する (S 6 5 0 6)。次に、サーバー装置のシステム制御部 6 4 2 1 は、音声合成処理部 6 4 2 3 に指令を出し、処理を開始させる。音声合成処理部 6 4 2 3 は、記憶部 6 4 2 2 から音声合成目的データを順次読み出しながら解析を行う。その際読み込んだデータが識別子で囲まれている場合はユーザ所望の音声キャラクタでない音素データベースを用い、そうでない場合はユーザ所望の音声キャラクタの音素データベースを用いる。

音声合成処理部 6 4 2 3 はデータを解析しながらそのデータに最も適する音素データを音素データベース保存部 6 4 2 5 から読み出して繋ぎ合わせ、合成音声データを作成し (S 6 5 0 7)、ユーザに送信する (S 6 5 0 8)。そしてユーザは合成音声データを端末装置に取りこみ再生することで端末装置からは合成音声データが出力される。

このようにこのシステムにおいては音声認証等での不正使用を防止するために合成目的データ中のある特定の文字列においてはユーザ所望のキャラクタ音声ではなく代替の音声で朗読することができる。

20 (実施の形態 2 6)

図 6 6 は本実施の形態の読上げシステムを示す構成図である。図 6 6 において要素 6 6 0 1 ~ 6 6 0 3 は (実施の形態 2 0) における図 4 9 の要素 4 9 0 1 ~ 4 9 0 3 と同一である。

この読上げシステムにおける携帯端末装置から出力される音声は音声認証等での不正使用を防止するために音声出力している文章の句点や句読点毎又は決められた文字数毎に音声合成の出力音であること示す音を強制的に挿入される。

図 6 7 は携帯端末装置 6 6 0 1 とサーバー装置 6 6 0 3 の詳細説明図である。要素 6 7 0 1 ~ 6 7 0 7 は (実施の形態 2 0) の図 4 9 の要素 4 9 0 1 ~ 4 9 0

7と同一である。またサーバー装置側の要素6721~6725は図49の要素4921~4925と同一である。

図68は読み上げシステムの動作フローチャートである。サーバー装置はまず、ユーザからの接続要求を待つ（S6801）。ユーザがユーザIDやパスワード等を用いて接続要求を行うと、サーバー装置のシステム制御部6721は接続要求を行ったユーザが正規ユーザかどうか判断する（S6802）。正規ユーザでなかった場合は接続拒否をユーザに通知する（S6803）。正規ユーザであった場合は接続許可および合成目的データ保存部6724にある合成目的データと音素データベース保存部6725にある音声キャラクタのリスト情報を送信する（S6804）。次にユーザが選択した所望の合成目的データと音声キャラクタデータのリクエストデータを受信すると（S6805）、サーバー装置のシステム制御部6721は該当の合成目的データを合成目的データベース6724から読み出し、サーバー装置の記憶部6722に記憶する。

次に、サーバー装置のシステム制御部6721は、音声合成処理部6723に指令を出し、処理を開始させる。音声合成処理部6723は、記憶部6722から音声合成目的データを順次読み出しながら解析を行う。まず音声合成処理部6723は合成処理済の文字数を格納する変数nを初期化し（S6806）、次に処理済みの文字数が識別音声挿入する文字数Tになったかを判断し（S6807）、そうである場合はnを0にして（S6808）、識別音データを合成音声データとして挿入する（S6809）。処理済みの文字数が識別音声挿入する文字数Tに達していない場合は、処理を行おうとしている文字が句読点または句点を示すデータかを判断し（S6810）、そうである場合は識別音データを合成音声データとして挿入する（S6809）。そうでない場合はユーザが選択した音声キャラクタの音素データベースから最適な音素を抽出する（S6811）。

そして識別音データ及び音素データベースから抽出された音素データは順次接続され、合成音声データが作成される（S6812）。音声合成処理部は合成処理を行った文字数の変数をカウントアップし（S6813）、合成目的データを全て音声合成処理したか判断し（S6814）、合成処理が終了している場合に

は、合成音声データをユーザに送信する（S6815）。そしてユーザは合成音声データを端末装置に取りこみ再生することで端末装置からは合成音声データが出力される。

- 5 このようにこのシステムにおいては音声認証等での不正使用を防止するために携帯端末装置から音声出力している文章の句点や句読点毎又は決められた文字数毎に音声合成の出力音であること示す音を強制的に挿入することができる。

（実施の形態27）

- 10 図69は読み上げシステムの概略説明図である。図69および図70において、端末装置本体6901は合成音データ入力手段とアンプ、スピーカ等を含んだ音声出力手段を備える。ここでの合成音データ入力手段とは、モデム等のネットワークインターフェースや光ディスク、磁気ディスク、メモリーカード等である記録媒体のデータ入力可能な記憶装置のインターフェースである。メモリーカードや光ディスク及び磁気ディスク等の記憶装置6902は合成音データ等を格納
- 15 し、端末装置本体6901とは脱着可能である。合成音声データ6903はサーバー手段から配信される。インターネット上のサーバー手段6904はユーザから指定された音声合成目的データと音声キャラクタの音素データベースを用いて音声合成を行い、合成音声データをユーザに配信する。

- 20 例えばユーザは端末装置本体6901を通じて、インターネット上のサーバー手段6904と通信し、サーバー手段6904に登録されている音声合成目的データを選択して、さらに選択した音声合成目的データの各データ範囲、例えば音声合成目的データが小説等であれば各登場人物のセリフ部分の音声合成に用いる音声キャラクタを選択する。サーバー手段6904は選択された音声キャラクタの音素データベースを用いて、音声合成目的データの音声合成を行い、その合成
- 25 音声データを通信手段を用いて、ユーザに配信する。ユーザはサーバー手段6904から配信された合成音声データを合成音データ入力手段を用いて端末装置本体6901に取り込み、再生することで所望の音声キャラクタでの合成音声を聴くことができる。

なお、サーバー手段6904は必ずしもインターネット上に無くてもよく、オフラインにてユーザからの要求を電話やFAX、郵便や人手で受け付け、合成音声データを光ディスクや磁気ディスク、メモリーカード等の記憶媒体に記録してユーザに配信してもよい。

- 5 図70は読み上げシステムのブロック図である。読み上げシステムは端末装置本体6901と記憶装置6902とサーバー手段6904とを備える。まずサーバー手段6904の各ブロックの説明を行う。

- サーバー手段6904において、サーバー制御部7000はサーバー手段全体の制御を行う。音声合成処理部7001は音声合成目的のデータの解析を行って、
10 各データに最適な音素データを抽出し連結する。音素データベース選択処理部7002は音声合成目的のデータを解析し、音声キャラクタを適用するデータ範囲を抽出して、各データ範囲の音声合成に用いる音素データベースを選択する。サーバー通信処理部7003は音声合成された合成音データをユーザに配信したり、ユーザとのインターフェースを行う。サーバー記憶部7004はサーバー手段全
15 体の制御を行うプログラムの保管やデータ処理の際の作業領域として用いられる。合成目的データ記録部7005は音声合成目的のデータを記録する。音素データベース記録部7006は音声キャラクタの音素データベースを記録する。音素データベースは、実在の人物の肉声をサンプリングし、そのサンプリングデータをデータベース化したものであり、出力される音声合成音の音色を決定する重要な
20 要素となる。

- 次に端末装置本体6901の各ブロックの説明を行う。端末装置本体6901において、端末制御部7007は装置内の各部とデータのやり取りを行い、装置全体の制御を行う。音声出力部7008は合成音データのフォーマット変換を行い、スピーカまたはヘッドフォンに出力する。合成音データ入力手段の一つである
25 記憶装置I/F部7009は記憶装置へのデータを読み書きする。端末記憶部7010は装置全体のプログラムの格納や様々な処理の作業領域として用いられる。操作部7011を通じユーザは装置に自分の指示を与える。表示部7012は装置の動作状態等をユーザに表示する。合成音データ入力手段の一つである端

末通信処理部7013はサーバー装置から送られてくる合成音データを受信したり、サーバー手段6904と端末装置本体6901のインターフェースを行う。電源部7014は装置に電源を供給する。

次に合成音データ入力手段の一つである記憶装置6902の各ブロックの説明を行う。端末装置I/F部7020は記憶装置I/F7009と共に端末装置本体6901とデータのやり取りを行う。合成音データ7021は記憶装置内部に記憶される。

次に本システムの詳細な動作説明を行う。

図71は本発明の読み上げシステムにおける動作フローチャートである。ユーザが端末装置本体6901の操作部7011を用いてサーバー手段6904との接続操作を行うと、端末通信処理部7013はサーバー手段6904と接続を行う。そしてユーザは小説等の音声合成目的データの選択要求を行う(S7101)。端末装置本体6901からの選択要求はサーバー通信処理部7003を通じ、サーバー手段6904に取り込まれ、サーバー制御部7000は端末装置本体6901からの音声合成目的データの要求を認識する(S7102)。

次にサーバー制御部7000は合成目的データ記録部内にある音声合成可能な合成目的データのリスト情報作成し、その情報を選択要求してきた端末装置本体6901に送る(S7103)。端末装置本体6901の端末制御部7007はサーバー手段6904から送られてきたリスト情報を認識して、その表示部7012に表示する(S7104)。そしてユーザは端末装置本体6901の操作部7011を用いて所望する音声合成目的データを選択決定する(S7105)。

次にサーバー制御部(7000)はユーザから選択決定された音声合成目的データを認識し(S7106)、該当のデータを合成目的データ記録部7005から読み出して、サーバー記憶部7004に記録する。次に音素データベース選択手段7002は音声合成目的のデータをサーバー記憶部7004から読み出しながら解析を

行い、各々の音素データベースを適用するデータの範囲を抽出する(S7107)。例えば音声合成目的のデータが小説のテキストデータの場合は、登場人物

は章毎や行毎であったりしても良く、その分け方は音声合成目的のデータ内容にも依存するので限定はしない。

次にサーバー制御部7000は音声合成処理部7001に処理を開始させる。

音声合成処理部7001はサーバー記憶部7004から音素データベース選択処

- 5 理部7002が処理したデータを順次読み出し、識別記号に基づき使用する音声キャラクタの音素データベースを選択し、同時に音声合成目的のデータを分析して、各データに最も適する音素データをサーバー記憶部7004または音素データベース記録部から読み出して、繋ぎ合わせ合成音データを作成する（S7113）。サーバー制御部7000は音声合成処理部7001が作成した合成音データ
- 10 ータをサーバー通信処理部7003を通じて

、ユーザに配信する（S7114）。サーバー手段6904から配信された合成音データは端末通信処理部7013を通じて、端末装置本体6901内の端末記憶部7010または記憶装置6902に記録される。そしてユーザが操作部7011を通じて再生の操作を行うと、合成音データが端末記憶

- 15 部7010または記憶装置から読み出されて音声出力部7008に渡される。音声出力部7008はデータのフォーマット変換を行い、合成音声をスピーカーまたはヘッドフォンに出力する（S7115）。

なお、端末通信処理部7013が端末装置本体6901に搭載されたものであるが、通信処理部を記憶装置6902に搭載し、ネットワーク上にあるサーバー

- 20 装置からデータをダウンロードして記憶装置に記憶するようにしても良い。

（実施の形態28）

図72は本実施の形態28の読上げシステムを示す構成図である。インターネット上のサーバー手段7201はユーザが音声合成を希望する音声合成目的データと音声キャラクタの音素データベースを用いて音声合成を行い、合成音声データ

25 ータをユーザに配信する。端末装置本体7202は合成音データ入力手段と、アンプ、スピーカ等を含んだ音声出力手段を備える。

ここでの合成音データ入力手段は、モデム等のネットワークインターフェース

や光ディスク、磁気ディスク、メモリーカード等である記録媒体のデータ入力が可能な記憶装置のインターフェースである。合成音声データ7203はサーバー手段7201から配信される。音声合成目的のデータ7204はユーザがサーバー手段7201に送信する。

- 5 ユーザはまず端末装置本体7202を通じて音声合成目的の文章が記述されたデータをサーバー手段7201に送信すると共に、自分が所望する音声キャラクタを選択する。サーバー手段7201は選択された音声キャラクタの音素データベースを用いてユーザから送信された音声合成目的データの音声合成を行い、合成音声データをインターネット経由でユーザに返信する。ユーザは端末装置本体
- 10 7202内にそのデータを取り込み、再生操作を行うことで端末装置本体7202からユーザが所望するキャラクタの音声でユーザが送信した文章データの合成音声が出来上がる。

- 端末装置本体7202とは脱着可能なメモリーカードや光ディスク及び磁気ディスク等の記憶装置7205は合成音データ等を格納する。なおユーザからの音
- 15 声合成の依頼やその受け付けはインターネット経由だけでなく、電話やファックス及び郵便や人手にて行われても良い。またサーバー手段7201からの合成音声データのユーザへの配信はインターネット経由だけではなく、光ディスクや磁気ディスク及びメモリーカード等の記憶媒体に合成音声データを記録し、それをユーザに配達してもよい。

- 20 図73は本実施の形態の読み上げシステムのブロック図である。図73において、読み上げシステムはサーバー手段7201と端末装置本体7202と記憶装置7203とを備える。まずサーバー手段7201の各ブロックの説明を行う。

- サーバー手段7201において、サーバー制御部7300はサーバー手段全体の制御を行う。音声合成処理部7301は音声合成目的のデータの解析を行って、
- 25 各データに最適な音素データを抽出し連結する。データ登録処理部7302はユーザから送られてくる音声合成目的のデータとユーザ情報を対応づけたデータを作成し、管理する。

サーバー通信処理部7303は音声合成された合成音データをユーザに配信し

たり、ユーザとのインターフェースを行う。サーバー記憶部7304はサーバー手段全体の制御を行うプログラムの保管やデータ処理の際の作業領域として用いられる。合成目的データ記録部7305は音声合成目的のデータを記録する。音素データベース記録部7306は音声キャラクタの音素データベースを記録する。

- 5 次に端末装置本体7202の各ブロックの説明を行う。端末装置本体7202において、端末制御部7307は装置内の各部とデータのやり取りを行い、装置全体の制御を行う。音声出力部7303は合成音データのフォーマット変換を行い、スピーカまたはヘッドフォンに出力する。合成音データ入力手段の一つである記憶装置I/F部7309は記憶装置へのデータを読み書きする。端末記憶部
- 10 7310は装置全体のプログラムの格納や様々な処理の作業領域として用いられる。操作部であり7311を通じユーザは装置に自分の指示を与える。表示部7312は装置の動作状態等をユーザに表示する。合成音データ入力手段の一つである端末通信処理部7313はサーバー装置から送られてくる合成音データを受信したり、サーバー手段7201と端末装置本体7202のインターフェースを
- 15 行う。電源部7314は装置に電源を供給する。データ入力処理部7315はユーザが音声合成目的のデータを入力する。

端末装置I/F部7320は記憶装置I/F部7309とともに端末装置本体7202とデータのやり取りを行う。合成音データ7321は記憶装置内部に記憶される。

- 20 図74は本実施の形態の読み上げシステムのフローチャートである。ユーザが端末装置本体7202の操作部7211を用いてサーバー手段7201との接続操作を行うと、端末通信処理部7213はサーバー手段7201と接続を行う。そしてユーザはサーバー手段7201に対して音声合成の要求を行う(S7401)。端末装置本体7202からの要求はサーバー通信処理部7303を通じ、
- 25 サーバー手段7201に取り込まれ、サーバー制御部7300はユーザからの音声合成要求を認識する(S7402)。次にサーバー制御部7300は音素データベース記録部7306内にある音声キャラクタのリスト情報を作成し、そのデータを端末装置本体7202に提供する(S7403)。

端末装置本体 7202 の端末制御部 7307 はサーバー手段 7201 から送られてきたリスト情報を認識して、その表示部 7312 に表示する (S7404)。そしてユーザは端末装置本体 7202 の操作部 7311 を用いて所望する音声キャラクターを選択決定する。またデータ入力処理部を用いて音声合成目的のデータを
5 端末装置本体 7202 に入力する。さらに操作部 7311 を用いてユーザの名前や住所、電話番号や E-MAIL アドレス、クレジット番号等のユーザ情報を入力する。そして端末制御部 7307 はこれらのデータをサーバー手段 7201 に伝える (S7405)。なおこのユーザ情報はユーザを特定でき、さらにサーバー手段 7201 がサービスに対する報酬を得る場合において、ユーザから料金を徴収する為に必要なデータであればよい。

次にサーバー制御部 7300 はユーザから選択決定された音声キャラクターと合成目的データ及びユーザ情報データを認識し (S7306)、合成目的データは合成目的データ記録部 7305 に記録し、ユーザ情報はサーバー記憶部 7304 に記録を行う。そしてデータ登録処理部 7302 は両データを対応づけるとともに、ユーザから受信した音声合成目的データのデータ量や音声キャラクター名等のデータもサーバー記憶部 7304 に記録する (S7407)。そしてこの対応付けたデータに基づき、サーバー手段 7201 がサービスに見合った報酬をユーザから徴収しても良い。

次にサーバー制御部 7300 は合成目的データ記憶部からユーザが合成依頼したデータを読み出してサーバー記憶部 7304 に記録し、音声合成処理部に処理を開始させる。順次読み出し、音声合成目的のデータを分析して、各データに最も適する音素データをサーバー記憶部 7304 または音素データベース記録部から読み出して、繋ぎ合わせ、合成音データを作成する (S7408)。サーバー制御部 7300 は音声合成処理部 7301 が作成した合成音データをサーバー通信処理部 7303 を通じて、ユーザに配信する (S7409)。

サーバー手段 7201 から配信された合成音データは端末通信処理部 7313 を通じて、端末装置本体内の端末記憶部 7310 または記憶装置に記録される。そしてユーザが操作部 7311 を通じて再生の操作を行うと、合成音データが端

末記憶部 7 3 1 0 または記憶装置から読み出されて音声出力部に渡される。音声出力部 7 3 0 8 はデータのフォーマット変換を行い、合成音声をスピーカーまたはヘッドフォンに出力する (S 7 4 1 0)。

10087421.030102

特許請求の範囲

1. 音素をデータ化した音素データベースに基づき、音声合成目的データを音声合成処理する音声合成処理部を有する端末装置本体と、

前記音声合成目的データと前記音素データベースとを保存し、前記端末装置本体と脱着可能な記憶装置とを備えた端末装置。

2. 前記音素データベースは人物の肉声をサンプリングしたサンプリングデータをデータベース化したものである、請求項1記載の端末装置。

3. 前記端末装置本体は、

音声合成処理部の出力を不要なノイズをカットしてスピーカに出力する音声出力処理部と、

ユーザが指示を与えるための操作部と、

ネットワークを介してサーバー装置と接続可能にする通信処理部と

をさらに有し、

前記音素データベースは人物の肉声をサンプリングしたサンプリングデータをデータベース化したものであり、

前記ユーザが前記操作部により前記音素データベースの種類や前記音声合成目的データを選択すると前記通信処理部は前記選択された前記音素データベースの前記種類と前記選択された音声合成目的データとの情報をサーバー装置へ送り、

音声合成処理部は前記サーバー装置より前記ネットワーク経由で送られた音素データベースと音声合成目的データを音声合成する、請求項1記載の端末装置。

4. 音声合成目的データと音素をデータ化した音素データとから音声合成処理

プログラムに基づいて音声合成を行う音声合成処理部と、

前記音声合成処理プログラムと前記音声合成目的データとを記憶する記憶部と、

記憶装置との間でデータの送受信を行うための記憶装置インタフェース部

5 と、

ネットワークに接続するための通信処理部と

を備え、前記記憶装置は、

前記音素データベースと、

前記端末装置本体との間でデータの送受信を行うための端末装置インタフ

10 ース部と

を備えた、端末装置。

5. 前記通信処理部は、ネットワーク上のサーバ装置から前記音声合成処理プログラムと前記音素データと前記音声合成目的データの少なくとも1つをダウンロード

15 ード可能である、請求項4記載の端末装置。

6. 音声合成目的データと音素をデータ化した音素データとから音声合成処理プログラムに基づいて音声合成を行う音声合成処理部と、

前記音声合成処理プログラムを記憶する記憶部と、

20 記憶装置との間でデータの送受信を行うための記憶装置インタフェース部と、

ネットワークに接続するための通信処理部と

を備え、前記記憶装置は、

前記音素データベースと、

25 前記音声合成目的データを格納する音声合成目的データ格納部と、

前記端末装置との間でデータの送受信を行うための端末装置インタフェース部と

を備えた端末装置。

10087421.030102

7. 音声合成目的データと音素をデータ化した音素データとから音声合成処理プログラムに基づいて音声合成を行う音声合成処理部と、

前記音声合成処理プログラムと前記音素データとを記憶する記憶部と、

5 記憶装置との間でデータの送受信を行うための記憶装置インタフェース部と、

ネットワークに接続するための通信処理部と

を備え、前記記憶装置は、

前記音声合成目的データを格納する音声合成目的データ格納部と、

10 前記記憶装置インターフェース部との間でデータの送受信を行うための端末装置インタフェース部と

を備えた、端末装置。

8. 音声合成目的データと音素をデータ化した音素データとから音声合成処理プログラムに基づいて音声合成を行う音声合成処理部と、

15 前記音声合成処理プログラムを記憶する記憶部と、

記憶装置との間でデータの送受信を行うための記憶装置インタフェース部と

を備え、前記記憶装置は、

20 前記音素データを格納する音素データベースと、

前記音声合成目的データを格納する音声合成目的データ格納部と、

前記記憶装置インターフェース部との間でデータの送受信を行うための端末装置インタフェース部と、

ネットワークに接続するための通信処理部と

25 を備えた、端末装置。

9. 前記通信処理部を介して前記音声合成処理プログラムと前記音素データと前記音声合成目的データとを前記ネットワーク上のサーバ装置からダウンロードし、

前記ダウンロードした音声合成目的データと音素データと音声合成処理プログラムとを前記端末装置インタフェース部と前記記憶装置インタフェース部とを介して前記記憶部に転送する記憶制御部をさらに備えた、請求項 8 記載の端末装置。

- 5 10. 前記音声合成処理プログラムを記憶する記憶部と、
記憶装置との間でデータの送受信を行うための記憶装置インタフェース部と、

ネットワークに接続するための通信処理部と
を備え、前記記憶装置は、

- 10 前記音声合成目的データと音素をデータ化した音素データとから前記音声合成処理プログラムに基づいて音声合成を行うと共に前記音声合成処理プログラムを格納する音声合成処理部と、

前記記憶装置インターフェース部との間でデータの送受信を行うための端末装置インタフェース部と

- 15 を備えた、携帯端末装置。

11. 前記記憶装置は前記音声合成目的データを格納する音声合成目的データベースと前記音素データを格納する音素データベースとの少なくとも 1 つをさらに備えた、請求項 10 記載の端末装置。

20

12. 前記端末装置本体と前記記憶装置のうちの 1 つはユーザがキャラクタ音声を登録する音声登録処理部をさらに有し、前記音声合成処理部は前記音声合成目的データを前記登録されたキャラクタ音声で音声合成する、請求項 1 記載の端末装置。

25

13. 前記端末装置本体と前記記憶装置のうちの 1 つは前記音声合成目的データをユーザが所望する言語に翻訳する翻訳処理部をさらに有し、前記音声合成処理部は前記翻訳された音声合成目的データを音声合成する、請求項 1 記載の端末

装置。

1 4. 前記端末装置本体と前記記憶装置のうちの 1 つはユーザがキャラクタ音声登録する音声登録処理部をさらに有し、前記音声合成処理部は前記翻訳された音声合成目的データを前記登録されたキャラクタ音声で音声合成する、請求項 1 3 記載の端末装置。

1 5. 前記端末装置本体は、ネットワークに接続するための通信処理部をさらに備え、前記通信処理部は、前記ネットワーク上のダウンロード選択処理部を備えた前記音声合成目的データを格納するサーバー装置から、前記音声合成目的データのうちのユーザが所望するの部分のみを前記端末装置本体と前記記憶装置のうちの 1 つにダウンロードする、請求項 1 記載の端末装置。

1 6. 前記音声合成目的データはテキストデータである、請求項 1 記載の端末装置。

1 7. 前記音声合成目的データは、楽譜データとテキストデータから成る楽曲データであり、前記音声合成処理部はユーザが前記音楽データベースと前記楽曲データを指定することで、ユーザが所望する特定キャラクタの音声で楽曲データを音声合成する、請求項 1 記載の端末装置。

1 8. 前記音声合成処理部は、前記音声合成目的データ中の特定の文字列をユーザ所望のキャラクタでの音声合成出力はせず、代替の音声で音声合成する、請求項 1 記載の端末装置。

1 9. 前記音声合成処理部は、前記音声合成目的データに音声合成の出力音であること示す音を挿入し音声合成する、請求項 1 記載の端末装置。

20. 前記端末装置本体と前記記憶装置の1つは前記音素データベースを選択する音素データベース選択手段を備え、前記音声合成処理部は前記音声合成目的データをその部分毎にユーザが所望するキャラクタ音声で音声合成する、請求項1記載の端末装置。

5

21. 前記端末装置本体と前記記憶装置のうちの1つは、

ユーザがキャラクタ音声を登録する音声登録処理部と、

前記音声合成目的データをそれに適用される前記音素データベースを選択する音素データベース選択手段と、

10

前記音声合成目的データと前記選択された音素データベースの識別信号とを格納する記憶部と

をさらに備え、前記音声合成処理部は前記識別記号を基に適用する音声キャラクタの音素データベースを判別する、請求項20記載の端末装置。

15

22. 前記音声合成目的データは楽譜と歌詞から成る楽曲データであり、

前記音声合成処理部はユーザが前記音素データベースと前記楽曲データを指定することで、ユーザが所望するキャラクタ音声で前記楽曲データを音声合成する、請求項20記載の端末装置。

20

23. 前記端末装置本体と前記記憶装置のうちの1つは、前記音声合成目的データに関する視覚表示を行う視覚表示処理手段をさらに備えた、請求項1記載の携帯端末装置。

25. 前記記憶装置は前記音声合成目的データを前記端末装置本体に入力し、

25

前記端末装置本体は前記音声合成処理部で音声合成された音声を出力するスピーカとイヤフォンのうちの少なくとも1つをさらに有する、請求項1記載の携帯端末装置。

1007421.030102

26. 前記記憶装置はメモリーカードと光ディスクと磁気ディスクとのうちの1つである、請求項1記載の携帯端末装置。

27. ネットワークとの通信機能を備えた制御部と、

- 5 音声合成目的データの各文字データに最適な音素データを抽出及び連結して、前記音素データを端末装置にある音声出力処理部に渡す音声合成処理部と、
前記音声合成目的データを保存するテキストデータ保存部と、

人物の肉声をサンプリングしたサンプリングデータをデータベース化した音素データベースを保存する音素データベース保存部と

- 10 を備え、

前記端末装置より音素データベースの種類や音声合成目的データを示す情報が送られると、前記音声合成処理部は前記送られた種類の音素データベースと前記送られた情報の示す音声合成目的データを音声合成し、

- 前記制御部は前記音声合成された音声合成データを前記ネットワーク経由
15 で前記端末装置へ送る、サーバー装置。

28. 楽譜データを保存する楽譜データ保存部と、

前記楽譜データ保存部から楽譜データを読み出し、さらに前記音声合成目的データを読み出し、前記楽譜データと音声合成された音を対にして繋ぎ合わせ、
20 前記端末装置にて再生可能なフォーマットに変換する楽曲合成処理部をさらに備えた、請求項27記載のサーバー装置。

29. 前記音素データベースを選択する音素データベース選択処理部をさらに備え、

- 25 前記音声合成処理部は前記音声合成目的データを部分毎にユーザが所望するキャラクタの音素データベースで音声合成し、

前記制御部は前記音声合成された音声合成目的データを前記ネットワーク経由で前記端末装置に送信し、

前記端末装置は前記受信した音声合成目的データを音声再生する、請求項
27記載のサーバー装置。

30. ユーザから送られてくる音声合成目的データとユーザ情報を対応付けて
5 管理するデータ登録処理部をさらに備え、

前記音声合成処理部は前記送られた音声合成目的データを音声合成し、

前記制御部は前記音声合成された合成音声データを前記ネットワーク経由
で前記端末装置に送信し、

前記端末装置は受信した前記音声合成目的データを音声再生する、請求項

10 27に記載のサーバー装置。

31. 音声合成処理部と音声合成目的データと音素データベースとを備えたネッ
トワーク上のサーバー装置と、

音声出力手段を有する端末装置と

15 を備え、

ユーザが前記端末装置を通じて音声合成目的データを選択すると、それに
従って前記音声合成処理部は指定された音素データベースから前記選択された音
声合成目的データを音声合成し、

前記サーバー装置は前記音声合成された音声合成目的データを前記ネット

20 ワーク経由で前記端末装置に配信し、

前記端末装置は前記音声合成目的データを音声再生する、読み上げシステ
ム。

32. 前記サーバー装置は、ユーザがキャラクタ音を登録する音声登録処理部を
25 さらに備え、前記音声合成処理部は前記登録されたキャラクタ音で前記音声合成
目的データを音声合成する、請求項31記載の読み上げシステム。

33. 前記サーバー装置は前記音声合成データをユーザの所望する言語に翻訳す

る翻訳処理部をさらに備え、

ユーザが前記端末装置を通じて音素データベースと前記言語の種類とを選択し、

前記音声合成処理部は前記翻訳された音声合成目的データを前記選択され

5 た音素データベースで音声合成する、請求項 3 1 記載の読み上げシステム。

3 4. 前記サーバー装置はユーザが所望するキャラクタ音声を登録する音声登録処理部をさらに備え、前記音声合成処理部は前記翻訳された音声合成目的データを前記キャラクタ音声で音声合成する、請求項 3 3 記載の読み上げシステム。

10

3 5. 前記音声合成目的データはテキストデータと楽譜データから成る楽曲データであり、前記音声合成処理部は前記音素データベースで前記楽曲データを音声合成する、請求項 3 1 記載の読み上げシステム。

15

3 6. 前記音声合成処理部は前記音声合成目的データ中の特定の文字列をユーザ所望のキャラクタでの音声合成せず、代替の音声で音声合成する、請求項 3 1 記載の読み上げシステム。

20

3 7. 前記音声合成処理部は前記音声合成目的データに、音声合成の出力音であることを示す音を挿入して音声合成する、請求項 3 1 記載の読み上げシステム。

3 8. 前記サーバー装置は前記合成された音声合成目的データのユーザが所望する部分のみを前記端末装置へダウンロード可能にするダウンロード選択処理部をさらに備えた、請求項 3 1 記載の読み上げシステム。

25

3 9. デジタル音声データをアナログに変換し、不要なノイズをカットしてスピーカに出力する音声出力処理部と、

ユーザが装置に指示を与えるための操作部と、

前記ユーザが前記操作部を通して選択した、人物の肉声をサンプリングしたサンプリングデータをデータベース化した音素データベースとテキストデータの情報をサーバー装置へ送るシステム制御部とを備え、

- 5 前記サーバー装置は前記指定された音素データベースより前記選択されたテキストデータを音声合成し、

前記音声出力処理部は前記サーバー装置からネットワーク経由で送られた前記音声合成されたテキストデータを出力する端末装置。

要約

携帯性に優れ、かつユーザが所望する声色で音声出力を行う携帯端末装置と読上げシステムを提供する。音声合成手段として実際の人物の肉声をサンプリングし、データベース化した音素データベースを使用して音声合成された音を出力する携帯端末装置において、ユーザは音声合成処理部や音素データベースや合成目的データが搭載された記憶装置を携帯端末装置に挿入し、再生することにより端末装置上で音声合成処理が行われ合成音声を聴くことができる。あるいは携帯端末装置をネットワーク上のサーバー装置と通信させ、自分が所望する音素データベースと合成目的データを選択することにより、ユーザは所望するキャラクタ音声で合成目的データである小説やニュース記事等の朗読を聴くことができる。